

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:32:24 ; Search time 212 Seconds
(without alignments)
2847.672 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMQNTCHRMSPHGRGCP...MREHREHQBPDDETDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_21.*
2: geneseqp1980s.*
3: geneseqp1990s.*
4: geneseqp2000s.*
5: geneseqp2001s.*
6: geneseqp2002s.*
7: geneseqp2003as.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	8	ADQ96675 Human rib
2	7500	100.0	1374	8	ADQ96675 Human rib
3	7500	100.0	1374	9	ADQ96675 Human rib
4	7486	99.8	1374	8	ADQ17464 Human rib
5	7486	99.8	1374	8	ADQ17464 Human rib
6	7483.5	99.8	1373	7	ADQ17386 Human rib
7	6426	85.7	1200	7	ADQ17386 Human rib
8	6062	80.8	1374	6	ADQ17386 Human rib
9	4071	54.3	769	4	ADQ17386 Human rib
10	2963	39.5	1327	4	ADQ17386 Human rib
11	2847	38.0	541	6	ADQ17386 Human rib
12	2519	33.6	486	9	ADQ17386 Human rib
13	2463	32.8	466	8	ADQ17386 Human rib
14	1961	26.1	378	4	ADQ17386 Human rib
15	1961	26.1	378	4	ADQ17386 Human rib
16	1961	26.1	378	4	ADQ17386 Human rib
17	1961	26.1	378	4	ADQ17386 Human rib
18	1330	17.6	263	7	ADQ17386 Human rib
19	1330	17.6	263	7	ADQ17386 Human rib
20	1295.5	17.3	301	4	ADQ17386 Human rib
21	1289	17.2	267	4	ADQ17386 Human rib
22	1122	15.0	277	4	ADQ17386 Human rib
23	890	11.9	412	8	ADQ17386 Human rib
24	890	11.9	412	8	ADQ17386 Human rib

25	890	11.9	412	9	ADQ17386 Human rib
26	702	9.4	148	5	ADQ17386 Human rib
27	496	6.6	115	2	ADQ17386 Human rib
28	496	6.6	115	3	ADQ17386 Human rib
29	472	6.3	97	5	ADQ17386 Human rib
30	340	4.5	1151	4	ADQ17386 Human rib
31	312.5	4.2	561	7	ADQ17386 Human rib
32	311	4.1	560	4	ADQ17386 Human rib
33	309	4.1	266	9	ADQ17386 Human rib
34	307.5	4.1	229	5	ADQ17386 Human rib
35	303.5	4.0	406	4	ADQ17386 Human rib
36	301.5	4.0	1663	8	ADQ17386 Human rib
37	298.5	4.0	228	5	ADQ17386 Human rib
38	298.5	4.0	228	8	ADQ17386 Human rib
39	298.5	4.0	228	8	ADQ17386 Human rib
40	298.5	4.0	228	8	ADQ17386 Human rib
41	298.5	4.0	242	5	ADQ17386 Human rib
42	297.5	4.0	800	8	ADQ17386 Human rib
43	285.5	3.8	225	9	ADQ17386 Human rib
44	285.5	3.8	1527	8	ADQ17386 Human rib
45	285.5	3.8	1636	8	ADQ17386 Human rib

ALIGNMENTS

RESULT 1	ADQ96675	ADQ96675 standard; protein; 1374 AA.
XX	ADQ96675;	
AC	ADQ96675;	
XX	23-SEP-2004 (first entry)	
DT	23-SEP-2004 (first entry)	
XX	Human ribonuclease III (RNase III) enzyme.	
DB	Human ribonuclease III (RNase III) enzyme.	
XX	Human; ribonuclease III; RNase III; research purpose; biological purpose;	
KW	clinical purpose; cellular interaction; enzyme.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	Region	1..220
FT	Region	/note="Proline rich region"
FT	Misc-difference	18
FT	Region	221..470
FT	Region	/note="Encoded by CGT"
FT	Region	949..1374
FT	Domain	/note="Serine-arginine rich region"
FT	Misc-difference	1211
FT	Region	/note="RNAse III domain
FT	Region	/note="Encoded by GCG"
FT	Region	1262..1269
FT	Region	/note="Alpha helix"
FT	Region	1282..1290
FT	Region	/note="Beta sheet"
FT	Region	1297..1303
FT	Region	/note="Beta sheet"
FT	Region	1308..1315
FT	Region	/note="Beta sheet"
FT	Region	1316..1336
FT	Region	/note="Alpha helix"
FT	Misc-difference	1345
FT	Misc-difference	/note="Encoded by GAA"
XX	US2004126867-A1.	
PN	US2004126867-A1.	
XX	01-JUL-2004.	
XX	01-JUL-2004.	
PD	01-JUL-2004.	
XX	09-FEB-2004; 2004US-00774974.	
XX	09-FEB-2004; 2004US-00774974.	
XX	06-JUL-2001; 2001US-00900425.	
PR	06-JUL-2001; 2001US-00900425.	
XX	06-JUL-2001; 2001US-00900425.	
XX	06-JUL-2001; 2001US-00900425.	

PA (CROO/) CROOKE S T.
 PA (WUHH/) WU H.
 XX
 PI Crooke ST, Wu H;
 XX
 DR MPI, 2004-516913/49.
 XX N-PSDB; ADQ96674.
 XX
 PT New isolated nucleic acid molecule encoding human RNase III, useful for
 PT research, biological, or clinical purposes, e.g. defining the roles of
 PT RNase III and the interaction of human RNase III and cellular RNA.
 XX
 PS Claim 3; SEQ ID NO 2; 31pp; English.
 XX
 CC The invention relates to human ribonuclease III (RNase III) and its
 CC corresponding nucleic acid sequence. The polynucleotide sequence of the
 CC invention is useful for research, biological and clinical purposes. It is
 CC useful in defining the roles of RNase III and the interaction of human
 CC RNase III and cellular RNA. The present sequence is the human RNase III
 CC enzyme.
 CC
 SQ Sequence 1374 AA;
 Query Match 100.0%; Score 7500; DB 8; Length 1374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MNOGNTCHMSFHPGRCGRGCGHGRSAPSFRQNTRLHPOQPVQYQYEPSPAS 60
 1 MNOGNTCHMSFHPGRCGRGCGHGRSAPSFRQNTRLHPOQPVQYQYEPSPAS 60
 61 TTFSSNPAENFLPRPDVFPFPPMPBSAOGPLPCPCIRPFPFNHQRHPFVPPCFPPM 120
 61 TTFSSNPAENFLPRPDVFPFPPMPBSAOGPLPCPCIRPFPFNHQRHPFVPPCFPPM 120
 121 PRPMPCNNPPVPGAPGCGTFRPMMPRPSMHPPPPPVQVQYQYPPGSHNFP 180
 121 PRPMPCNNPPVPGAPGCGTFRPMMPRPSMHPPPPPVQVQYQYPPGSHNFP 180
 121 PRPMPCNNPPVPGAPGCGTFRPMMPRPSMHPPPPPVQVQYQYPPGSHNFP 180
 181 SFNSFONNSSPFLPSANNSSPHFRLLPYPLPKASERRSPERLKHVDHRRHDSGR 240
 181 SFNSFONNSSPFLPSANNSSPHFRLLPYPLPKASERRSPERLKHVDHRRHDSGR 240
 181 SFNSFONNSSPFLPSANNSSPHFRLLPYPLPKASERRSPERLKHVDHRRHDSGR 240
 241 GERHSLDRERGRSPDRRRODSRYSDYDRGRTSPRHSYERSRERERHRRHNRSS 300
 241 GERHSLDRERGRSPDRRRODSRYSDYDRGRTSPRHSYERSRERERHRRHNRSS 300
 241 GERHSLDRERGRSPDRRRODSRYSDYDRGRTSPRHSYERSRERERHRRHNRSS 300
 301 PELERSYKKEKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSMAPLEIVNHRSPSRK 360
 301 PELERSYKKEKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSMAPLEIVNHRSPSRK 360
 301 PELERSYKKEKRSGRSYGLSVVPEPAGCTPELPGELIKNTDSMAPLEIVNHRSPSRK 360
 361 KQARWEEEDDRMSDNOSSGDKNYTSIKEKEPEETMPDKNEEELKPVWIRCHSEN 420
 361 KQARWEEEDDRMSDNOSSGDKNYTSIKEKEPEETMPDKNEEELKPVWIRCHSEN 420
 421 YVSSDPMDOVGSTVGTSLRLDLYDKFEELSGRQEKAAARPWEPPKTKLDEDLSS 480
 421 YVSSDPMDOVGSTVGTSLRLDLYDKFEELSGRQEKAAARPWEPPKTKLDEDLSS 480
 481 SSESCESDSDSTSSSDSEVFVIAEIKRKAHPRLHDELMYNDPGQNDGPLCKCSA 540
 481 SSESCESDSDSTSSSDSEVFVIAEIKRKAHPRLHDELMYNDPGQNDGPLCKCSA 540
 541 KARRTGIRHSIYGEBAIKPCRPMTNAGRLFYRTITVSPPTNFLDRPTVIEYDHEYT 600
 541 KARRTGIRHSIYGEBAIKPCRPMTNAGRLFYRTITVSPPTNFLDRPTVIEYDHEYT 600
 601 PEGFSMFAPAPLTNIPDLCKVIRENIDYTIHFLEEMNPENFCVAGLEFSLFRDLIELY 660
 601 PEGFSMFAPAPLTNIPDLCKVIRENIDYTIHFLEEMNPENFCVAGLEFSLFRDLIELY 660
 661 DNNLKGPLFEDSPCCPRFHEMPRFVRLPDGSKVLSHQIILYLIRCSKALVPEEETA 720
 661 DNNLKGPLFEDSPCCPRFHEMPRFVRLPDGSKVLSHQIILYLIRCSKALVPEEETA 720

DB 661 DNNLKGPLFEDSPCCPRFHEMPRFVRLPDGSKVLSHQIILYLIRCSKALVPEEETA 720
 QY 721 NMLQWEELEWQCYAECECKGMITVNTPGTKSSVRIQDLDRQFNPDVTPIIVHRCIRPA 780
 DB 721 NMLQWEELEWQCYAECECKGMITVNTPGTKSSVRIQDLDRQFNPDVTPIIVHRCIRPA 780
 QY 781 QLSVAGDDPOYQCLMKSIVYKRLHLANSFKVQKTDQOKLAOREALOKIRQKNTMREYLV 840
 DB 781 QLSVAGDDPOYQCLMKSIVYKRLHLANSFKVQKTDQOKLAOREALOKIRQKNTMREYLV 840
 QY 841 ELSSQGFMTKTRISDVCOHAWMLPVLTHHIRHQCIMLMDKLGITFORCLLQALMTNP 900
 DB 841 ELSSQGFMTKTRISDVCOHAWMLPVLTHHIRHQCIMLMDKLGITFORCLLQALMTNP 900
 QY 901 SHLHFGNPDHARNSLNSCGIRQPKYGRKVNHNMRKKGINTLINIMSRLGODDPTSS 960
 DB 901 SHLHFGNPDHARNSLNSCGIRQPKYGRKVNHNMRKKGINTLINIMSRLGODDPTSS 960
 QY 961 RINHNRELEFLGDVAVEFLTSVHLVYLPSPSEGGIATYRTAIYONQHLAMLAKKLELD 1020
 DB 961 RINHNRELEFLGDVAVEFLTSVHLVYLPSPSEGGIATYRTAIYONQHLAMLAKKLELD 1020
 QY 1021 FMLYAHGPDLCESDLRHMANCFEALIGAVYLEGSLSEAKOLFGRLLFNDPDLREVMIN 1080
 DB 1021 FMLYAHGPDLCESDLRHMANCFEALIGAVYLEGSLSEAKOLFGRLLFNDPDLREVMIN 1080
 QY 1081 YELHPLQLOEPNTDRLQLETSFVLQKLEFEBAIGVIFTHVRLARAFTLRTVGFNHLL 1140
 DB 1081 YELHPLQLOEPNTDRLQLETSFVLQKLEFEBAIGVIFTHVRLARAFTLRTVGFNHLL 1140
 QY 1141 GHNQMEFLGDSIMOLVATEYLFIHPDHHHEGHTLRLSSLVNNTQAAVAEELGQGEYA 1200
 DB 1141 GHNQMEFLGDSIMOLVATEYLFIHPDHHHEGHTLRLSSLVNNTQAAVAEELGQGEYA 1200
 QY 1201 ITNDTKRPVGRRTKTLADLESFIALAYTDKOLEVHTFMNVCFEPRLKEFLINDAMD 1260
 DB 1201 ITNDTKRPVGRRTKTLADLESFIALAYTDKOLEVHTFMNVCFEPRLKEFLINDAMD 1260
 QY 1261 PKSQLOCCCLTLRTGKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCGGSPISQQA 1320
 DB 1261 PKSQLOCCCLTLRTGKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCGGSPISQQA 1320
 QY 1321 EMGAAMDALAEKNFQMAHQKFIKRYKQELKEMRWEEHQBREDDEDIKK 1374
 DB 1321 EMGAAMDALAEKNFQMAHQKFIKRYKQELKEMRWEEHQBREDDEDIKK 1374

RESULT 2
 ADR68727
 ID ADR68727 standard; protein; 1374 AA.
 XX
 AC ADR68727;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human Ribonuclease III, RNase III.
 XX
 KW Human; enzyme; Ribonuclease III; RNase III; RNA interference;
 KW gene silencing; double stranded RNA; pre-rRNA processing; RNA processing;
 KW RNA expression; RNA splicing; RNA translocation.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..220
 FT /label = Proline_rich_domain
 FT /note = "Claimed in claim 8"
 FT Domain 221..470
 FT /label = Serine Arginine_rich_domain
 FT /note = "Claimed in claim 8"
 FT Domain 499..1374
 FT /label = RNase_III_domain
 FT /note = "Claimed in claim 8"

FT Misc-difference 1211 /note= "Encoded by GCG"
 FT Misc-difference 1345 /note= "Encoded by GAA"
 FT US2004175828-A1.
 PD 09-SEP-2004.
 XX 22-MAR-2004; 2004US-00805919.
 XX 06-JUL-2001; 2001US-00900425.
 PA (CROO/) CROOKE S T.
 PA (WUHH/) WU H.
 PI Crooke ST, Wu H;
 XX WPI; 2004-667111/65.
 DR
 XX Modulating RNA interference, processing, expression, splicing and
 XX translocation in cell or tissue, by contacting cell or tissue with human
 PT RNAse III or oligomeric compound targeted to nucleic acid encoding human
 PT RNAse III.
 XX
 XX Claim 7; SEQ ID NO 2; 32pp; English.
 CC The invention relates to modulating RNA interference (gene silencing),
 CC RNA processing (e.g. of pre-rRNA), RNA expression, RNA splicing, or RNA
 CC translocation in a cell or tissue, by contacting the cell or tissue with
 CC a modulator effective to cause the modulation by at least 50% as compared
 CC to control, where modulator is human RNAse III (ribonuclease III, which
 CC cleaves double stranded RNA) polypeptide or an oligomeric compound (an
 CC antisense oligonucleotide) targeted to the nucleic acid encoding human
 CC RNAse III. The oligomeric compound is targeted to a 3'-untranslated
 CC region (3'UTR), a 5'-untranslated region (5'UTR) or a coding region of a
 CC nucleic acid molecule encoding human RNAse III polypeptide, where the
 CC oligomeric compound inhibits the expression of human RNAse III
 CC polypeptide by at least 50%. In the method above, the RNA is RNA, snRNA,
 CC snRNA or miRNA, or precursors of the above. The RNA is processed to form
 CC one or more 12S pre-RNA and 28S rRNA fragments. The methods are useful
 CC for modulating RNA interference in a cell or tissue, modulating
 CC processing of an RNA in a cell or tissue, modulating RNA expression in a
 CC cell or tissue, modulating RNA splicing in a cell or tissue, and
 CC modulating RNA translocation in a cell or tissue. The methods are useful
 CC for research, biological and clinical purposes. The methods are useful in
 CC defining the roles of RNAse III and interaction of human RNAse III and
 CC cellular RNA. The present sequence represents human RNAseIII.
 XX
 XX Sequence 1374 AA:
 90
 Query Match 100.0%; Score 7500; DB 8; Length 1374;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 GERHRSILDRERGRSDRRRDRSRYSDDRGRTSPRRHSYRSRERERHRHRNRSS 300
 QY 301 PSLSRYKKEYRSGSYGLSVPEPAGCTPELPGRIKNTGSMAPLEIIVNHRSPSRK 360
 DB 301 PSLSRYKKEYRSGSYGLSVPEPAGCTPELPGRIKNTGSMAPLEIIVNHRSPSRK 360
 QY 361 KRARWEEKDRWSDNQSCKDKNTYSIKKEPEETPDKNSEEEBELKPMWIRCTHSEN 420
 DB 361 KRARWEEKDRWSDNQSCKDKNTYSIKKEPEETPDKNSEEEBELKPMWIRCTHSEN 420
 QY 421 YSSDPMDVGDSTVYVGTIRLDLDYDKFEEELGSRQEKAKARPPWEPKTKLDEDESS 480
 DB 421 YSSDPMDVGDSTVYVGTIRLDLDYDKFEEELGSRQEKAKARPPWEPKTKLDEDESS 480
 QY 481 SESECEDESDTCSSSSDSEVPDVAIEIRKKAHPRLDELMYNDPGOMDGPLCKCA 540
 DB 481 SESECEDESDTCSSSSDSEVPDVAIEIRKKAHPRLDELMYNDPGOMDGPLCKCA 540
 QY 541 KARRTGIRHSIYGEBAIKPCRPMTNAGRLPHYRITVSPTNPLTDRPTVIEYDHEYT 600
 DB 541 KARRTGIRHSIYGEBAIKPCRPMTNAGRLPHYRITVSPTNPLTDRPTVIEYDHEYT 600
 QY 601 FEGFSMFAPAPLTNIPLCKVIRPNIDYTHFIEEMWPEVCYKGLBESLFLPRDILEY 660
 DB 601 FEGFSMFAPAPLTNIPLCKVIRPNIDYTHFIEEMWPEVCYKGLBESLFLPRDILEY 660
 QY 661 DNMLKGPLFEDSPCCPRFHPMPFRPLPDGKEVLSHGQITLYLRSCKLVPEEETA 720
 DB 661 DNMLKGPLFEDSPCCPRFHPMPFRPLPDGKEVLSHGQITLYLRSCKLVPEEETA 720
 QY 721 NMLQWEELEWQYAECECKMIVTNPETKPSYRIDLDREQNPDPVETPIIVHFGIRPA 780
 DB 721 NMLQWEELEWQYAECECKMIVTNPETKPSYRIDLDREQNPDPVETPIIVHFGIRPA 780
 QY 781 QLSYAGDPYOYQKLMSYVYKLRHLANSYVKQTDYKQKLAQREBALQIKQKTMREVTY 840
 DB 781 QLSYAGDPYOYQKLMSYVYKLRHLANSYVKQTDYKQKLAQREBALQIKQKTMREVTY 840
 QY 841 ELSOGFMWTKGIRSPVCOHAMMLPVLTNHIRVHOCIMHLDKIGYFQDRCLQLAMTHP 900
 DB 841 ELSOGFMWTKGIRSPVCOHAMMLPVLTNHIRVHOCIMHLDKIGYFQDRCLQLAMTHP 900
 QY 901 SHHLNFGMNPDAHARSLNSCGIRQPKYGDRAKVNHMRRKKGINTLINIMSRIGQDDPTS 960
 DB 901 SHHLNFGMNPDAHARSLNSCGIRQPKYGDRAKVNHMRRKKGINTLINIMSRIGQDDPTS 960
 QY 961 RINNHERLEFLGDAVVEFLTSVHLVYLPSPLEEGGIATYRTAIVONQHLAMLAKKLEDP 1020
 DB 961 RINNHERLEFLGDAVVEFLTSVHLVYLPSPLEEGGIATYRTAIVONQHLAMLAKKLEDP 1020
 QY 1021 EMLYAHGPDLCRESOLRHNAMACFEALIGAVVLEGSLEBAKOLFGLLPNDLDREVMIN 1080
 DB 1021 EMLYAHGPDLCRESOLRHNAMACFEALIGAVVLEGSLEBAKOLFGLLPNDLDREVMIN 1080
 QY 1081 YELHPLQLOEPRTDQLETSPVLQKLEFEFEALIGVIFTHVLLARAFTLRVGFNHLTL 1140
 DB 1081 YELHPLQLOEPRTDQLETSPVLQKLEFEFEALIGVIFTHVLLARAFTLRVGFNHLTL 1140
 QY 1141 GHNQMEFLGDSIMOLVATEYLFIHPDDHBEGLTLRSSLVNNRTQAKVAEELGQEYA 1200
 DB 1141 GHNQMEFLGDSIMOLVATEYLFIHPDDHBEGLTLRSSLVNNRTQAKVAEELGQEYA 1200
 QY 1201 ITNDKTKRPVGRRTKTLADLESFIAALYTDQLELVHFMVNCPEPRPKETELINDQND 1260
 DB 1201 ITNDKTKRPVGRRTKTLADLESFIAALYTDQLELVHFMVNCPEPRPKETELINDQND 1260
 QY 1261 PKSQLOQCCLTARTGKEPEDIPLYKTLQTVGSHARTYVAAYFGEKIRIGCGKSPSIOQA 1320
 DB 1261 PKSQLOQCCLTARTGKEPEDIPLYKTLQTVGSHARTYVAAYFGEKIRIGCGKSPSIOQA 1320
 QY 1321 EWGAAMDALEKYNPOMAHQKRFIGRKYRQELKEMWEREHOEDBEDTEDIKK 1374

Db 1321 ENGAAMDLEKXNPQMAHQKRFGRKRYQBLKEMKWEREHQERBDETEDIK 1374

RESULT 3

ID AEB47411 standard; protein; 1374 AA.

AC AEB47411;

XX 22-SEP-2005 (first entry)

DE Human ribonuclease III (RNase III) protein, SEQ ID NO: 2.

KM Ribonuclease; gene silencing; enzyme.

OS Homo sapiens.

XX

Key Location/Qualifiers

FT Region 1..220

FT /note= "RNase III proline-rich region"

FT Misc-difference 18

FT /note= "Encoded by CGT"

FT Region 221..470

FT /note= "RNase III serine-arginine rich region"

FT Domain 949..1374

FT /note= "RNase III domain"

FT Misc-difference 1211

FT /note= "Encoded by GCG"

FT Region 1262..1269

FT /note= "Alpha helix"

FT Region 1282..1290

FT /note= "Beta sheet"

FT Region 1297..1303

FT /note= "Beta sheet"

FT Region 1308..1315

FT /note= "Beta sheet"

FT Region 1318..1336

FT /note= "Alpha helix"

FT Misc-difference 1345

FT /note= "Alpha helix"

FT /note= "Encoded by GAA"

XX

US2005159384-A1.

XX

21-JUL-2005.

XX

02-DEC-2004; 2004US-00001993.

XX

06-JUL-2001; 2001US-00900425.

XX

20-FEB-2002; 2002US-00079185.

XX

(ISIS-) ISIS PHARM INC.

XX

Crooke ST;

XX

WPI; 2005-512270/52.

XX

N-PSDB; AEB47410.

XX

GENBANK; AAF80558.

XX

Use of RNase III for eliciting modification of a selected RNA target,

XX

promoting gene silencing of a gene, inhibiting the expression of a gene,

XX

promoting inhibition of expression of a gene, or eliciting modification

XX

of an RNA target.

XX

Claim 19; SEQ ID NO 2; 32pp; English.

XX

The present invention relates to ribonuclease III (RNase III) nucleic

XX

acids and their encoding proteins. RNase III is an endoribonuclease that

XX

cleaves double stranded RNA. The invention is useful for eliciting

XX

modification of a selected RNA target in a cell, promoting gene silencing

XX

of a gene, inhibiting the expression of a gene, promoting inhibition of

XX

expression of a gene and eliciting modification of a RNA target in a

XX

cell. The present sequence is human ribonuclease III (RNase III) protein.

SQL Sequence 1374 AA:

Query Match 100.0%; Score 7500; DB 9; Length 1374;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMQNTGCHRMSPHRCGCPGRGGHGAAPASAPSPQNLRLHPQOPVQYGEPPSAS 60

DB 1 MMQNTGCHRMSPHRCGCPGRGGHGAAPASAPSPQNLRLHPQOPVQYGEPPSAS 60

QY 61 TTFNSPAPNPLPPDPVPPPPPPPSAQGLPPCPPIPPPPNQMRHPVPPPCFPM 120

DB 61 TTFNSPAPNPLPPDPVPPPPPPPSAQGLPPCPPIPPPPNQMRHPVPPPCFPM 120

QY 121 PPMPCPNPPVPYGAAPGQGTFFPMMPSPMHPPPPPVPMQVYVYPPGYSHNFP 180

DB 121 PPMPCPNPPVPYGAAPGQGTFFPMMPSPMHPPPPPVPMQVYVYPPGYSHNFP 180

QY 181 SPSNQQNNPSSFLPSANSSSPHFHLLPPYPLPKAPSERSEPERLKHVDHHRHSHGR 240

DB 181 SPSNQQNNPSSFLPSANSSSPHFHLLPPYPLPKAPSERSEPERLKHVDHHRHSHGR 240

QY 241 GERHSLDRERGRSPDRRQDSRYRSDYDRGRTSPRHSYRSEREERHRHNR 300

DB 241 GERHSLDRERGRSPDRRQDSRYRSDYDRGRTSPRHSYRSEREERHRHNR 300

QY 301 PSLERSYKKEYKRSGRSYGLSVPEPACCTPELPEIINKTDSWAPPLEIVNHSPSREK 360

DB 301 PSLERSYKKEYKRSGRSYGLSVPEPACCTPELPEIINKTDSWAPPLEIVNHSPSREK 360

QY 361 KRAKWEKDRMSDQSSGCKDNYTSIKEKEPEEPMPONEEBELKPVNIRCHSN 420

DB 361 KRAKWEKDRMSDQSSGCKDNYTSIKEKEPEEPMPONEEBELKPVNIRCHSN 420

QY 421 YSSDPMDQVGDSTVGTSTRLDLYDKFEEBELSGROEAKAARPMPEPKTLDDELSS 480

DB 421 YSSDPMDQVGDSTVGTSTRLDLYDKFEEBELSGROEAKAARPMPEPKTLDDELSS 480

QY 481 SESECESEDSDSTCSSSDSEVEFVLAIEIKRKAHPDRDLDELWYNDPGMDGPLICSA 540

DB 481 SESECESEDSDSTCSSSDSEVEFVLAIEIKRKAHPDRDLDELWYNDPGMDGPLICSA 540

QY 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAGRLPHYATITSPPNPLTDRRTVLEYDHEEY 600

DB 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAGRLPHYATITSPPNPLTDRRTVLEYDHEEY 600

QY 601 FEGFMPAHPALTNTPLCKVIRFNIDYTHFIEEMPEFVCYKGLFSLFPRDILEY 660

DB 601 FEGFMPAHPALTNTPLCKVIRFNIDYTHFIEEMPEFVCYKGLFSLFPRDILEY 660

QY 661 DNMLKGPLFEDSPCCPRFHPFRFVRLPDGKEVLSMHQILLYLRCSKALVPEEEL 720

DB 661 DNMLKGPLFEDSPCCPRFHPFRFVRLPDGKEVLSMHQILLYLRCSKALVPEEEL 720

QY 721 NMLQWEELEWQYAECECKMITYNPGTKSSVRILQDLDEQNPVITPPIIVHFEIRA 780

DB 721 NMLQWEELEWQYAECECKMITYNPGTKSSVRILQDLDEQNPVITPPIIVHFEIRA 780

QY 781 QLSYAGDPQYCKLWMSYVLRHLANSPKVKOTDKLQREBALOKIRQKTMRELEV 840

DB 781 QLSYAGDPQYCKLWMSYVLRHLANSPKVKOTDKLQREBALOKIRQKTMRELEV 840

QY 841 ELSQGFWMKTGIRSDVCGHMMPLVLTTHIRYHOCIMLIDLKIGYTFQDRCLQLAMTHP 900

DB 841 ELSQGFWMKTGIRSDVCGHMMPLVLTTHIRYHOCIMLIDLKIGYTFQDRCLQLAMTHP 900

QY 901 SHHLNFGANNPDAARNSLSCGIRQPKYGRKYNHMMKKGINTLINISRIGQDDPT 960

DB 901 SHHLNFGANNPDAARNSLSCGIRQPKYGRKYNHMMKKGINTLINISRIGQDDPT 960

QY 961 RINNHERLEFGDAVEFLTSYHLVLPSPLEGGATARTAYVONQHLAMLAKKLEDP 1020

DB 961 RINNHERLEFGDAVEFLTSYHLVLPSPLEGGATARTAYVONQHLAMLAKKLEDP 1020

QY 1021 FMUYAHGPDLCRESDDL RHAMANCFEALIGAVYLEGSLBEAKOLFGLLFNDPDLREVMIN 1080
DB 1021 FMUYAHGPDLCRESDDL RHAMANCFEALIGAVYLEGSLBEAKOLFGLLFNDPDLREVMIN 1080
QY 1081 YPLHPIQLOEPNTDROLITSPVLQKLTETEEBAIGVIFTHVRLLAFAFLIRTVGFNHLTL 1140
DB 1081 YPLHPIQLOEPNTDROLITSPVLQKLTETEEBAIGVIFTHVRLLAFAFLIRTVGFNHLTL 1140
QY 1141 GNNORHEFLGDSIMQVATEVYLIFHRPDHHEGLTLIRSSLVNNRQAVAEELGQOEYA 1200
DB 1141 GNNORHEFLGDSIMQVATEVYLIFHRPDHHEGLTLIRSSLVNNRQAVAEELGQOEYA 1200
QY 1201 INNDTKRPVGLRTKTLADLLESFIALYTDKOLEVHTFMVNCFFPRLKEFLINDMDND 1260
DB 1201 INNDTKRPVGLRTKTLADLLESFIALYTDKOLEVHTFMVNCFFPRLKEFLINDMDND 1260
QY 1261 PXSULOQCCLTLRTEGKEPDIPLYKTLQTVGSPSHATTVAVYFKGERIGCGKGPSIQQA 1320
DB 1261 PXSULOQCCLTLRTEGKEPDIPLYKTLQTVGSPSHATTVAVYFKGERIGCGKGPSIQQA 1320
QY 1321 EHGAMDALEKYNFPQMAHOKRPIGRKYNOELKEMWEEERHOREDEDEDIKK 1374
DB 1321 EHGAMDALEKYNFPQMAHOKRPIGRKYNOELKEMWEEERHOREDEDEDIKK 1374

RESULT 4

ADQ17464
ID ADQ17464 standard; protein; 1374 AA.

ADQ17464;

DT 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated protein - SEQ ID 281.

KX soft tissue sarcoma; cytosolic; gene therapy; vaccine; screening; human.

OS Homo sapiens.

FN WO2004048938-A2.

PD 10-JUN-2004.

PF 26-NOV-2003; 2003WO-US038193.

PR 26-NOV-2002; 2002US-0429739P.

PA (PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Ginsburg WM, Zlotnick A;

XX WPI; 2004-441208/41.

DR Early detection of soft tissue sarcoma comprises determining expression

PT of a gene in a first soft tissue sample and a normal soft tissue sample

PT and comparing the gene expression, also useful in treating soft tissue

PT sarcoma.

XX Example 2; SEQ ID NO 281; 210bp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma

CC which comprises obtaining a first soft tissue sample from an individual

CC and a normal soft tissue sample from the same or different individual,

CC determining the expression of a gene in both samples and comparing the

CC expression of the gene in both soft tissue samples, where a higher level

CC of protein expression in the first soft tissue sample indicates the

CC presence of soft tissue sarcoma. The method of the invention has

CC cytostatic applications and may be useful for detecting soft tissue

CC sarcoma, possibly via gene therapy or vaccine production. The nucleic

CC acid sequences may be useful in diagnostic and screening applications.

CC The current sequence is that of a human soft tissue sarcoma-upregulated

CC protein of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.

XX Sequence 1374 AA;

Query Match 99.8%; Score 7486; DB 8; Length 1374;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 1372; Conservative 0; Matched 2; Indels 0; Gaps 0;

QY 1 MMQAGTCHMSFHPGRCPRGRGAGAPSPAPSPFQNLRLHPQPPVQYQYEPSPAS 60
DB 1 MMQAGTCHMSFHPGRCPRGRGAGAPSPAPSPFQNLRLHPQPPVQYQYEPSPAS 60
QY 61 TTFSNSPABNPLPRBDVPFPFPMPBSAQGLPCCPIRPPFNHOMRHPFPVPCFP 120
DB 61 TTFSNSPABNPLPRBDVPFPFPMPBSAQGLPCCPIRPPFNHOMRHPFPVPCFP 120
QY 121 PPMPCPNPVPVGPAGAPGCGTFFPMMPPPSMHPFPVPPVPPVPPVPPVPPVPP 180
DB 121 PPMPCPNPVPVGPAGAPGCGTFFPMMPPPSMHPFPVPPVPPVPPVPPVPPVPP 180
QY 181 SFNSFQNNPSSFLPSANNSSPHFRLLPYPLPKAPSEERSPERLKHVDHRHDSHGR 240
DB 181 SFNSFQNNPSSFLPSANNSSPHFRLLPYPLPKAPSEERSPERLKHVDHRHDSHGR 240
QY 241 GERHRLDRERERGRSPDRRRODSRYSDYDRGRTSPRHSYERSREERHRHRDRRS 300
DB 241 GERHRLDRERERGRSPDRRRODSRYSDYDRGRTSPRHSYERSREERHRHRDRRS 300
QY 301 PSLERSYKKEYSKSGSYGLSVPEPAGCTPELPGIINOTDSWAPPLIVNHRSSREK 360
DB 301 PSLERSYKKEYSKSGSYGLSVPEPAGCTPELPGIINOTDSWAPPLIVNHRSSREK 360
QY 361 KARWEEKDRMSDNOSGDKNYTISIKKEBPETPDKNBEEBELKPWIRCTHSEN 420
DB 361 KARWEEKDRMSDNOSGDKNYTISIKKEBPETPDKNBEEBELKPWIRCTHSEN 420
QY 421 YSSDPMDOVGDSTVVGTSRLDLYDKFEEELSGRQEKAAAPWPEPKTLDLELSS 480
DB 421 YSSDPMDOVGDSTVVGTSRLDLYDKFEEELSGRQEKAAAPWPEPKTLDLELSS 480
QY 481 SSESCESDSDSTCSSSDSEVPDVAIEIKKKAHPRLHDELMTYNDPGQNDGFLCKGA 540
DB 481 SSESCESDSDSTCSSSDSEVPDVAIEIKKKAHPRLHDELMTYNDPGQNDGFLCKGA 540
QY 541 KARTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTVTSPTNPLFDRPTVLEEDHEYI 600
DB 541 KARTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTVTSPTNPLFDRPTVLEEDHEYI 600
QY 601 FEGFSMFAAPLTNIPDLCKVIRFNIDYTHIFTEEMAPENFCVGLFSLFRDILELY 660
DB 601 FEGFSMFAAPLTNIPDLCKVIRFNIDYTHIFTEEMAPENFCVGLFSLFRDILELY 660
QY 661 DNLKGPFLFEDSPCCPRFHPMPFVRLPDGKEVLSHGQILLYLRCSKALVPEEEL 720
DB 661 DNLKGPFLFEDSPCCPRFHPMPFVRLPDGKEVLSHGQILLYLRCSKALVPEEEL 720
QY 721 NMLQWEELMOKYAEBCCKMIVTNPCTKSSVRIIDLDEQFNPDVITPPIYHFGIRPA 780
DB 721 NMLQWEELMOKYAEBCCKMIVTNPCTKSSVRIIDLDEQFNPDVITPPIYHFGIRPA 780
QY 781 QLSYAGDPQYOKLMSYVQLRHLANSPKYKQTDOKLQAREEALOKIRQKNTMREEVY 840
DB 781 QLSYAGDPQYOKLMSYVQLRHLANSPKYKQTDOKLQAREEALOKIRQKNTMREEVY 840
QY 841 ELSSQGFMTGTIRSDVCOHAMMLPVITTHIRYHQCIMHLDKLIGYTFORCLQLMTTP 900
DB 841 ELSSQGFMTGTIRSDVCOHAMMLPVITTHIRYHQCIMHLDKLIGYTFORCLQLMTTP 900
QY 901 SHHLNGMPPDHARNLSNCGTROPKYGRKYNHMMRKKGINTLNNISRLGQDDPTPS 960
DB 901 SHHLNGMPPDHARNLSNCGTROPKYGRKYNHMMRKKGINTLNNISRLGQDDPTPS 960
QY 961 RINHNRELEFLGDAVVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNHMLAMLAKKLEDP 1020

|||||
 Db RINHERLEFLGDAVEFLTSVHLVLPSPLEGGATYRTAIVONOHMLAKKLELDP 1020
 QY 961 RINHERLEFLGDAVEFLTSVHLVLPSPLEGGATYRTAIVONOHMLAKKLELDP 1020
 QY 1021 FWLYHGPDLCESDLRHMANCFEALIGAVYLESGLESAKOLPGRLEFNDDLEWMLN 1080
 Db 1021 FWLYHGPDLCESDLRHMANCFEALIGAVYLESGLESAKOLPGRLEFNDDLEWMLN 1080
 QY 1081 YPLHPLQOEPTNDRLIETSPVLQCLTEFEEAIGVIFTHVRLARAFPLRTVGFNHLTL 1140
 Db 1081 YPLHPLQOEPTNDRLIETSPVLQCLTEFEEAIGVIFTHVRLARAFPLRTVGFNHLTL 1140
 QY 1141 GHNORMEFSDSIMQVATEYLFIFPHDHGHLTLIRSSLVNNRTOAVAEELGQOEYA 1200
 Db 1141 GHNORMEFSDSIMQVATEYLFIFPHDHGHLTLIRSSLVNNRTOAVAEELGQOEYA 1200
 QY 1201 IINDTKRVRGRTKLADLLESFIALATYTDKOLEVHTFMNVCFPRLKEFILNDMDND 1260
 Db 1201 IINDTKRVRGRTKLADLLESFIALATYTDKOLEVHTFMNVCFPRLKEFILNDMDND 1260
 QY 1261 PKSOIQCCCLTTRTEGKEPDIPLYKTLQTVGSHARTYTVAVYFKGERIGCGKGPISIOQA 1320
 Db 1261 PKSOIQCCCLTTRTEGKEPDIPLYKTLQTVGSHARTYTVAVYFKGERIGCGKGPISIOQA 1320
 QY 1321 EWGAAMDALCKNPFQMAHQKRFIGKTYQELKEMWREHOREPDETDIKK 1374
 Db 1321 EWGAAMDALCKNPFQMAHQKRFIGKTYQELKEMWREHOREPDETDIKK 1374
 QY 1374 EWGAAMDALCKNPFQMAHQKRFIGKTYQELKEMWREHOREPDETDIKK 1374
 Db 1374 EWGAAMDALCKNPFQMAHQKRFIGKTYQELKEMWREHOREPDETDIKK 1374

RESULT 5
 ADY17386
 ID ADY17386 standard; protein; 1374 AA.
 AC ADY17386;
 XX
 DT 05-MAY-2005 (first entry)
 XX
 DE PRO polypeptide SEQ ID NO 3192.
 XX
 KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
 XX Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
 KW Antichyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
 KW Vincicid; Gastrointestinal-Gen.; Antipsoriatic; Antiaesthetic;
 KW Antiallergic; ds; gene; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FN WO2005016962-A2.
 XX
 PD 24-FEB-2005.
 XX
 PF 11-AUG-2004; 2004WO-US026249.
 XX
 PR 11-AUG-2003; 2003US-0493546P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
 XX
 DR WPI, 2005-182350/19.
 XX
 PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
 XX treating an immune related disorder, e.g. systemic lupus erythematosus,
 CC rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
 CC
 PS Claim 8; SEQ ID NO 3192; 158bp; English.
 XX
 CC The invention relates to an isolated nucleic acid encoding a PRO
 CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
 CC composition, and method are useful for diagnosing and treating an immune
 CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
 CC arthritis. The present sequence represents a DNA encoding a PRO
 CC polypeptide.
 XX

Sequence 1374 AA:
 Query Match 99.8%; Score 7486; DB 9; Length 1374;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MMQNTGCHRMSPHPRGCGPRGGHGAAPASAPSPONILRLHPQOPVQYQYEPSPAS 60
 Db 1 MMQNTGCHRMSPHPRGCGPRGGHGAAPASAPSPONILRLHPQOPVQYQYEPSPAS 60
 QY 61 TTFSSPAENFLPRPDDVPPPPPPSAOGLPCCPIRPPPPNHQMRHPPPPPCFP 120
 Db 61 TTFSSPAENFLPRPDDVPPPPPPSAOGLPCCPIRPPPPNHQMRHPPPPPCFP 120
 QY 121 PPPMPCPNPPVPYGAAPGQGTFFMMPPSPMHPPPPPVPMQOVVQYPPGYSHNFP 180
 Db 121 PPPMPCPNPPVPYGAAPGQGTFFMMPPSPMHPPPPPVPMQOVVQYPPGYSHNFP 180
 QY 181 SPNSFQNNPSSFLPSANNSSPHFRHLPPYLPKAPSERSRSPERLKHVDHRRHDSHGR 240
 Db 181 SPNSFQNNPSSFLPSANNSSPHFRHLPPYLPKAPSERSRSPERLKHVDHRRHDSHGR 240
 QY 241 GERHSLDRRERGRSPDRRRODSRYSDYDRGRTSPRRHSYRSRERERHRRHNRSS 300
 Db 241 GERHSLDRRERGRSPDRRRODSRYSDYDRGRTSPRRHSYRSRERERHRRHNRSS 300
 QY 301 PSLERSYKKEYKRSRSGSYGLSVPPAGCTPELPEIINKTDSMAPLEIVNHRSPSRBK 360
 Db 301 PSLERSYKKEYKRSRSGSYGLSVPPAGCTPELPEIINKTDSMAPLEIVNHRSPSRBK 360
 QY 361 KQARWEEBKDRWSDNOSSGDKNTYSIKEKEEETMPDNNEEBEELKPVVIRCHSHSN 420
 Db 361 KQARWEEBKDRWSDNOSSGDKNTYSIKEKEEETMPDNNEEBEELKPVVIRCHSHSN 420
 QY 421 YSSDPMOVGSDTVTGTSTRLDLYDKFEELGSGROEKAKAPRPMPEPKTLDLEDJSS 480
 Db 421 YSSDPMOVGSDTVTGTSTRLDLYDKFEELGSGROEKAKAPRPMPEPKTLDLEDJSS 480
 QY 481 SSECESDESDTSCSSSDSEVFDVLAIEIKKXAHPRDLDELWYNDPQMGNDGPLCKCSA 540
 Db 481 SSECESDESDTSCSSSDSEVFDVLAIEIKKXAHPRDLDELWYNDPQMGNDGPLCKCSA 540
 QY 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAGRLHYHAIYVSPPNFLTRDTVIEYDHEXI 600
 Db 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAGRLHYHAIYVSPPNFLTRDTVIEYDHEXI 600
 QY 601 FEGSPMPAARPLTNIPLCKVIRFNTDYYTHFTEEMMPENFCYKGLFELFPRDILLEY 660
 Db 601 FEGSPMPAARPLTNIPLCKVIRFNTDYYTHFTEEMMPENFCYKGLFELFPRDILLEY 660
 QY 661 DNNLKGPLFEDSPCCPRFHEMPREVRFLPDGKEVLSMHOILLVLRCSKALVPEEBIA 720
 Db 661 DNNLKGPLFEDSPCCPRFHEMPREVRFLPDGKEVLSMHOILLVLRCSKALVPEEBIA 720
 QY 721 NMLQWEELEWQKYAECEKMIYTNPGTKRSSYRIQDLDREONPDVITPPIIVHGFIRA 780
 Db 721 NMLQWEELEWQKYAECEKMIYTNPGTKRSSYRIQDLDREONPDVITPPIIVHGFIRA 780
 QY 781 QLSYAGDPQYOKLMTSYVYLRLHLSNSPKVKOTDKLAQREBALOKIKQKTMRESEVY 840
 Db 781 QLSYAGDPQYOKLMTSYVYLRLHLSNSPKVKOTDKLAQREBALOKIKQKTMRESEVY 840
 QY 841 ELSOSGFMTGIRSDVCOHAMMLPVLTHIRYHOCIMLMDKLGTYFODRCLQLAMTHP 900
 Db 841 ELSOSGFMTGIRSDVCOHAMMLPVLTHIRYHOCIMLMDKLGTYFODRCLQLAMTHP 900
 QY 901 SHHLNFGANPDHARNSLNSGIRDPKYGDKYHNMNMKKGINTLINISRLGODPTPS 960
 Db 901 SHHLNFGANPDHARNSLNSGIRDPKYGDKYHNMNMKKGINTLINISRLGODPTPS 960
 QY 961 RINHERLEFLGDAVEFLTSVHLVLPSPLEGGATYRTAIVONOHMLAKKLELDP 1020
 Db 961 RINHERLEFLGDAVEFLTSVHLVLPSPLEGGATYRTAIVONOHMLAKKLELDP 1020

QY 1021 FMLYAGPDLCSRESDLRHMANCFEALIGAVLEGSLEBEAKQULFGRLLFNDPDLREVMAN 1080
 CC |||||
 CC target in a cell comprises contacting an RNA-like polynucleotide-RNA
 CC target duplex with a polypeptide comprising an RNase III domain. Also
 CC included are promoting gene silencing in a cell, inhibiting the
 CC expression of a gene in a cell comprising employing the method of cited
 CC above, promoting inhibition of expression of a gene, a hybrid RNase III
 CC (comprising at least one domain from a human RNase III and at least one
 CC domain from an RNase III of an organism other than human) and a cell
 CC having enhanced RNase III activity over an activity exhibited by a second
 CC cell (where the second cell is not enriched with respect to the amount or
 CC activity of RNase III polypeptide). The method is useful for eliciting a
 CC modification of a selected RNA target in a cell, and for promoting
 CC inhibition of expression of a gene in a cell. Compositions comprising
 CC RNase III polypeptides or polynucleotides are useful for research,
 CC biological and clinical purposes. The polynucleotides are may be used in
 CC defining the roles of RNase III and the interaction of human RNase III
 CC and cellular RNA. Host cells can be used for the production of human
 CC RNase III and for identifying agents, which increase or decrease levels
 CC of expression or activity of human RNase III in the cell. The present
 CC sequence represents human RNase III (a double stranded RNase)
 CC
 QY 1141 GHNORHEPLGDSIMOLVATEYLFIHPDHEGHLTLRSLVNNRTOAKVAEELGMOEYA 1200
 CC |||||
 CC
 QY 1201 ITNDTKRPVGLRTKTLADLLESFIALYTDKLEVTFFMVCFFPRLKEFLINDMND 1260
 CC |||||
 CC
 QY 1201 ITNDTKRPVGLRTKTLADLLESFIALYTDKLEVTFFMVCFFPRLKEFLINDMND 1260
 CC |||||
 CC
 QY 1261 PXSQLOCCCLTTRTEGKEPDIPLYKTLQTVGSPSHANTYVAVYFKGERRIGCGSPSIQQA 1320
 CC |||||
 CC
 QY 1261 PXSQLOCCCLTTRTEGKEPDIPLYKTLQTVGSPSHANTYVAVYFKGERRIGCGSPSIQQA 1320
 CC |||||
 CC
 QY 1321 EKGAAWDALAEKYNPQMAHOKRFIRGRKYQELKEMWREPHOREPDEDEDIKK 1374
 CC |||||
 CC
 QY 1321 EKGAAWDALAEKYNPQMAHOKRFIRGRKYQELKEMWREPHOREPDEDEDIKK 1374
 CC |||||
 CC

RESULT 6
 ABU63361
 ID ABU63361 standard; protein; 1373 AA.
 XX
 AC ABU63361;
 XX
 DT 18-SEP-2003 (first entry)
 XX
 DE Human double stranded RNase, RNase III.
 XX
 KW Human; enzyme; RNase III; double stranded RNase; RNA target;
 KW gene silencing.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 18 /note= "Encoded by CGT"
 FT Misc-difference 1211 /note= "Encoded by GCG"
 FT Misc-difference 1345 /note= "Encoded by GAA"
 FT
 PN US2003044941-A1.
 PD
 PD 06-MAR-2003.
 PF 20-FEB-2002; 2002US-00079185.
 PR 06-JUN-1996; 96US-00659440.
 PR 06-JUN-1997; 97US-00870608.
 PR 07-JAN-2000; 2000US-00479783.
 PR 06-JUL-2001; 2001US-00900425.
 XX
 XX (CROO/) CROOKE S T.
 PA
 PI Crooke ST;
 XX
 DR WPI; 2003-521756/49.
 DR N-PSDB; ACD27526.

eliciting a modification of a selected RNA target in a cell, useful for
 promoting inhibition of gene expression in a cell, comprises contacting
 an RNA-like polynucleotide-RNA target duplex with a polypeptide having an
 RNase III domain.

claim 37; Fig 1; 33pp; English.

XX
 CC The invention relates to eliciting a modification of a selected RNA
 CC target in a cell comprises contacting an RNA-like polynucleotide-RNA
 CC target duplex with a polypeptide comprising an RNase III domain. Also
 CC included are promoting gene silencing in a cell, inhibiting the
 CC expression of a gene in a cell comprising employing the method of cited
 CC above, promoting inhibition of expression of a gene, a hybrid RNase III
 CC (comprising at least one domain from a human RNase III and at least one
 CC domain from an RNase III of an organism other than human) and a cell
 CC having enhanced RNase III activity over an activity exhibited by a second
 CC cell (where the second cell is not enriched with respect to the amount or
 CC activity of RNase III polypeptide). The method is useful for eliciting a
 CC modification of a selected RNA target in a cell, and for promoting
 CC inhibition of expression of a gene in a cell. Compositions comprising
 CC RNase III polypeptides or polynucleotides are useful for research,
 CC biological and clinical purposes. The polynucleotides are may be used in
 CC defining the roles of RNase III and the interaction of human RNase III
 CC and cellular RNA. Host cells can be used for the production of human
 CC RNase III and for identifying agents, which increase or decrease levels
 CC of expression or activity of human RNase III in the cell. The present
 CC sequence represents human RNase III (a double stranded RNase)
 CC
 QY 1 MMGNTCHRMSPHFGGCGCRGCGHAGRPASBFRQNLRLHPQOPVQYQYEPSPAPS 60
 CC |||||
 CC
 QY 1 MMGNTCHRMSPHFGGCGCRGCGHAGRPASBFRQNLRLHPQOPVQYQYEPSPAPS 60
 CC |||||
 CC
 QY 61 TTFNSPAPNFPLPRDPVFPFPMPSPSAQSPLPCCPIRPFPPNHNMRPFPPECFPM 120
 CC |||||
 CC
 QY 61 TTFNSPAPNFPLPRDPVFPFPMPSPSAQSPLPCCPIRPFPPNHNMRPFPPECFPM 120
 CC |||||
 CC
 QY 121 PPMPCPNPVPVGAAPGQTFPPMMPSPSPMPRPPVPMQGVNYQYRPGYSHNPPPP 180
 CC |||||
 CC
 QY 121 PPMPCPNPVPVGAAPGQTFPPMMPSPSPMPRPPVPMQGVNYQYRPGYSHNPPPP 180
 CC |||||
 CC
 QY 181 SFNSFQNNPSSFLPSANNSSPHFRLLPPYPLPKASERSSPRLKGYDHRDRSHGR 240
 CC |||||
 CC
 QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTPSRHSYERSRERERHRRDRRS 300
 CC |||||
 CC
 QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTPSRHSYERSRERERHRRDRRS 300
 CC |||||
 CC
 QY 301 PSLERSYKKEYKRGSGSYGLSVVPEBAGCTPELPGBIINKTDSWAPPLEIVNHRSPSREK 360
 CC |||||
 CC
 QY 301 PSLERSYKKEYKRGSGSYGLSVVPEBAGCTPELPGBIINKTDSWAPPLEIVNHRSPSREK 360
 CC |||||
 CC
 QY 361 KRAREEEDKRWSDNOSGDKKNTYSIKKEPEETMPDKNEEBEELKVMIRCHSEN 420
 CC |||||
 CC
 QY 361 KRAREEEDKRWSDNOSGDKKNTYSIKKEPEETMPDKNEEBEELKVMIRCHSEN 420
 CC |||||
 CC
 QY 421 YSSDPMDOVGDSTVGTSLRLDLYKKFEEELGSRQEKAAAPRMEPPRTKJDEPLESS 480
 CC |||||
 CC
 QY 421 YSSDPMDOVGDSTVGTSLRLDLYKKFEEELGSRQEKAAAPRMEPPRTKJDEPLESS 480
 CC |||||
 CC
 QY 481 SESSECSDESDTSCSSSDSEVPDVAIEIKRKAHPRLDELWYNDPGQNDGSLCKSA 540
 CC |||||
 CC
 QY 481 SESSECSDESDTSCSSSDSEVPDVAIEIKRKAHPRLDELWYNDPGQNDGSLCKSA 540
 CC |||||
 CC
 QY 541 KARTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTTVPSPNTPLDRPVIYDYDHEYI 600
 CC |||||
 CC
 QY 541 KARTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTTVPSPNTPLDRPVIYDYDHEYI 600
 CC |||||
 CC
 QY 601 FSGFSMAFAHPLNTPLCVIRINIDYTHFIEEMBPENPCVGLFSLFRDITELY 660
 CC |||||
 CC
 QY 601 FSGFSMAFAHPLNTPLCVIRINIDYTHFIEEMBPENPCVGLFSLFRDITELY 660
 CC |||||
 CC
 QY 661 DWNLKGPLFEDSPCCPRFHFMRPVRPLPDGKEVLSMHQILLYLRCSKALVPEEEL 720
 CC |||||
 CC

Query Match 99.8%; Score 7483.5; DB 7; Length 1373;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1373; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

661	DNLKGPFLPEDESPCCPRPFHMPKRVRLPDGGAEVLSMHQILLYLAFCSKALNPDEBIA	720
AY	721 NMLQWEELEWOKYAECKGMIVTNPETKPSVVRIDQLDREQFNDVITPPIIVHFGIRPA	780
DB	721 NMLQWEELEWOKYAECKGMIVTNPETKPSVVRIDQLDREQFNDVITPPIIVHFGIRPA	780
OY	781 OLSYAGDPQOYQKLWMSYVKLRLHLANSRKVQTDQYQKLAQREBALQKLRQNTWREBVTY	840
DB	781 QLSYAGDPQOYQKLWMSYVKLRLHLANSRKVQTDQYQKLAQREBALQKLRQNTWREBVTY	840
OY	841 ELSGSGFPKTGISRSVCOHAMMLPYLTNHRIRHOCIMHLDKLGTFPDRCILQLAMTHP	900
DB	841 ELSGSGFPKTGISRSVCOHAMMLPYLTNHRIRHOCIMHLDKLGTFPDRCILQLAMTHP	900
OY	901 SHHLNFGNPPDHARNSLSNCGIRQPKYDGRKVVHNMKKGKINTLINIMSRIGQDDPTPS	960
DB	901 SHHLNFGNPPDHARNSLSNCGIRQPKYDGRKVVHNMKKGKINTLINIMSRIGQDDPTPS	960
OY	961 RINRNERLEFGDAVVEPLTSHLYLPPSLBEGGLATYRPAIVONOHMLAMLAKKLEIDP	1020
DB	961 RINRNERLEFGDAVVEPLTSHLYLPPSLBEGGLATYRPAIVONOHMLAMLAKKLEIDP	1020
OY	1021 FMLYAHGPDLCRESDLRHHMANCFEALGAVYLEGSLBEAKOLFRLFPNDPDLREVMIN	1080
DB	1021 FMLYAHGPDLCRESDLRHHMANCFEALGAVYLEGSLBEAKOLFRLFPNDPDLREVMIN	1080
OY	1081 YPLHPLQLOEPNTDROLIETSPVLQKLTPEBEALGVIETHVRLIARAFLLRTVGFNNHLL	1140
DB	1081 YPLHPLQLOEPNTDROLIETSPVLQKLTPEBEALGVIETHVRLIARAFLLRTVGFNNHLL	1140
OY	1141 GHNOQMEFLGDSIMQVATETYLFIHPDHHEGHLTLFSSLVNRRTOAKVAEBELGMQRYA	1200
DB	1141 GHNOQMEFLGDSIMQVATETYLFIHPDHHEGHLTLFSSLVNRRTOAKVAEBELGMQRYA	1200
OY	1201 ITNDTKRPVGLRTTLADLLESFPAALYTBDILEYVTTFMNVCFEPFLKKEFILNQDND	1260
DB	1201 ITNDTKRPVGLRTTLADLLESFPAALYTBDILEYVTTFMNVCFEPFLKKEFILNQDND	1260
OY	1261 PKSQLOQCCLTLRTGSKSPDIPLYKTLQTVGFSHARTYTVAVYFKGBRIGCGKGPSIOQA	1320
DB	1261 PKSQLOQCCLTLRTGSKSPDIPLYKTLQTVGFSHARTYTVAVYFKGBRIGCGKGPSIOQA	1320
OY	1321 EMGAAMDALAEKYNFPOMAHQKRFGRKTRQELKEMRWERHOERPDDETDIKK	1374
DB	1321 EMGAAMDALAEKYNFPOMAHQKRFGRKTRQELKEMRWERHOERPDDETDIKK	1374
RESULT 7		
AEAZ0195	AEAZ0195 standard; protein; 1200 AA.	
AC	AEA20195;	
DT	11-AUG-2005 (first entry)	
DE	Novel human polypeptide SEQ ID NO 889.	
KW	vulnery; CNS-gen.; gene therapy; diagnostic; forensic; mapping;	
KM	DNA purification; protein purification; osteoarthritis; antiarthritic;	
KM	osteoporotic; musculoskeletal disease; osteoporosis; endocrine disease;	
KM	peridontal disease; antiinflammatory; mouth disease; burns; injury;	
KM	peripheral neuropathy; Alzheimer disease; neuroprotective; nootropic;	
KM	degeneration; parkinson disease; antiparkinsonian; neurological disease;	
KM	cerebrovascular ischemia; cerebroprotective; vasotropic;	
KM	cardiovascular disease; autoimmune disease; immunosuppressive;	
KM	immune disorder; viral infection; virucide; infection; cancer;	
KM	cytostatic; neoplasm.	
OS	Homo sapiens.	
XX		
FN	WO2005049806-A2.	

PD		02-JUN-2005.	
PE		11-MAR-2004; 2004WO-US007412.	
PX		14-MAR-2003; 2003US-00389559.	
XX		(NUVE-) NOVELO INC.	
PA		Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;	
PI		Ghosh M, Xue A, Abundi V, Zhao Q, wang D, Goodrich R, Chen R;	
PI		Weinman T, Weng G, Boyle B;	
XX		WPI; 2005-417730/42.	
DR		N-PsDB; AEA19628.	
XX		New polynucleotide encoding a polypeptide with biological activity,	
PT		useful for treating a disease or disorder, e.g. osteoarthritis, burns,	
PR		CNS and peripheral disease, stroke, autoimmune disorders, viral	
PT		infection, or cancer.	
PS		Claim 20; SEQ ID NO 889; 500bp; English.	
PX		The invention describes a new isolated polynucleotide (I) encoding a	
CC		polypeptide with biological activity comprising: a nucleotide sequence of	
CC		SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes	
CC		to the sequence of (i) under stringent hybridization conditions; or a	
CC		nucleotide sequence having greater than 99% sequence identity with the	
CC		sequence of (I). Also described are: a(n) (expression) vector comprising	
CC		(I); a host cell genetically engineered to comprise (I) operatively,	
CC		associated with a regulatory sequence that modulates expression of the	
CC		polynucleotide in the host cell; an isolated polypeptide comprising a	
CC		sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide	
CC		is: a polypeptide encoded by (I); or a polypeptide encoded by a	
CC		polynucleotide hybridizing under stringent conditions with any one of SEQ	
CC		ID NOS: 1-567; a composition comprising the polypeptide of (3) and a	
CC		carrier; an antibody directed against the polypeptide of (3); a method	
CC		for detecting (I) in a sample; a method for detecting the polypeptide of	
CC		(3) in a sample; a method for identifying a compound that binds to the	
CC		polypeptide of (3); a method of producing the polypeptide of (3); and a	
CC		collection of polynucleotides, where the collection comprising of at	
CC		least one of SEQ ID NOS: 1-567. (I) is a polynucleotide comprising any of	
CC		the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological	
CC		activity, which comprises any of the amino acid sequence of SEQ ID NOS:	
CC		568-1134. All sequences are fully defined in the specification. The	
CC		sequences and methods are useful in diagnostics, forensic, and gene	
CC		mapping, in identifying of mutations responsible for genetic disorders or	
CC		types of traits, in assessing biodiversity, and for producing many other	
CC		types of data and products dependent on DNA and amino acid sequences. The	
CC		composition and method are useful for treating a disease or disorder,	
CC		e.g. osteoporosis, osteoarthritis, periodontal disease, burns, CNS and	
CC		peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,	
CC		autoimmune disorders, viral infection, or cancer. This is the amino acid	
CC		sequence of a novel polypeptide of the invention.	
XX		Sequence 1200 AA:	
QY	Query Match	85.7%; Score 6426; DB 9; Length 1200;	
Db	Best Local Similarity	94.1%; Pred. No. 0;	
	Matches 1193; Conservative	1; Mismatches 6; Indels 68; Gaps 1	
QY	107	MHHPPVPCPPMPMPCNNPVPCAPAGQGCFPMMPSPMHPPVPVMOQNY	166
Db	1	MHHPPVPCPPMPMPCNNPVPCAPAGQGFPMMPSPMHPPVPVMOQXY	60
QY	167	OYPGYSHNHPSPSFGONNPGSLFSANSSGPHFRHLPPYLPAKPSRRSRPERLX	226
Db	61	OYPPYSHNHPSPSFGONNPGSLFSANSSGPHFRHLPPYLPAKPSRRSRPERLX	120
QY	227	HYDDHRHDSHGGRGERHSILDRERGSPPRRRDSFYRSDDYDGRTPSRHRSYERE	286
Db	121	HYDDHRHDSHGGRGERHSILDRERGSPPRRRDSFYRSDDYDGRTPSRHRSYERS--	178
QY	287	RERENHRNRDNRRSGLSERYKYKKYKRGSGYGLSVPEPACGPTELGEIIIXNTDSMAP	346

```

Db      179 ----- 178
Qy      347 PLEIVNRSPSREKKAARWEEKDRMSDNOSGKDKNYTS IKKEPEETWPDQNEEBEE 406
Db      179 -----SRPSREKKAKARWEEKDRMSDNOSGDKNYTS IKKEPEETWPDQNEEBEE 232
Qy      407 LKKPWIRCTHSENNYSSDPMQVGS TVVGTSLRLDLYDKFEELGSROEKAKAAPPM 466
Db      233 LKKPWVVRCTHSENNYSSDPMQVGS TVVGTSLRLDLYDKFEELGSROEKAKAAPPM 292
Qy      467 EPPKTKLDDDLSSSSSECESEDSTCSSSDESEVDVLAETKKAHPDRLHDELMTND 526
Db      293 EPPKTKLDDDLSSSSSECESEDSTCSSSDESEVDVLAETKKAHPDRLHDELMTND 352
Qy      527 PQOMNDGPKCSAKARRTGIRHSIYPGEEAIKPCRPMTNNAGRLPHYRTVSPNPLT 586
Db      353 PQOMNDGPKCSAKARRTGIRHSIYPGEEAIKPCRPMTNNAGRLPHYRTVSPNPLT 412
Qy      587 DRPTVLEYDDHEXIFEGFSMFAHAPLTNIPLCKVIRFNIDYTIHPIEMMPENFCVKGLE 646
Db      413 DRPTVLEYDDHEXIFEGFSMFAHAPLTNIPLCKVIRFNIDYTIHPIEMMPENFCVKGLE 472
Qy      647 LPSLFLFRDILBYDMNLKGPLFEDSPCCPRPFHMPFRVRLPDGKKEVLSMHQILLYL 706
Db      473 LPSLFLFRDILBYDMNLKGPLFEDSPCCPRPFHMPFRVRLPDGKKEVLSMHQILLYL 532
Qy      707 LKCSKALVBEERANMLQWEELEWOKYAECKGMITNPGTKSSVRIDQLDREOQNPV 766
Db      533 LKCSKALVBEERANMLQWEELEWOKYAECKGMITNPGTKSSVRIDQLDREOQNPV 592
Qy      767 ITFPIIVHFGIRPAQSYAGDPQYQKLMKSYVGLRHLANS PKVKOTDQKLAQREBALQ 826
Db      593 ITFPIIVHFGIRPAQSYAGDPQYQKLMKSYVGLRHLANS PKVKOTDQKLAQREBALQ 652
Qy      827 KTRQKNTMRREVTVELSSQGFMTGIRSDVCOHAMMLPVLTNHIRYHOCIMHLDKILGYT 886
Db      653 KTRQKNTMRREVTVELSSQGFMTGIRSDVCOHAMMLPVLTNHIRYHOCIMHLDKILGYT 712
Qy      887 FODRCILQIOLAMTHPSHNLNPGNMPDHAARSLNSCGIROPKYGRKYNHNMKKGINTLI 946
Db      713 FODRCILQIOLAMTHPSHNLNPGNMPDHAARSLNSCGIROPKYGRKYNHNMKKGINTLI 772
Qy      947 NIMSRLEGODPTPSRINNHRELEFLGDAVVEFLTSVHLYYLFPSSLEGGIATRYTAIVON 1006
Db      773 NIMSRLEGODPTPSRINNHRELEFLGDAVVEFLTSVHLYYLFPSSLEGGIATRYTAIVON 832
Qy      1007 QHILAMIAKKLELDPPFMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEBAKQLFGR 1066
Db      833 QHILAMIAKKLELDPPFMYAHGPDLCRESDLRHAMANCFEALIGAVYLEGSLEBAKQLFGR 892
Qy      1067 LFPNDPDLAEWMLNLYLHLQLOEPNTDROLIETSVYLOKITEFEBAIGVIFHVLAKR 1126
Db      893 LFPNDPDLAEWMLNLYLHLQLOEPNTDROLIETSVYLOKITEFEBAIGVIFHVLAKR 952
Qy      1127 AFLTLRTVGFNHLTLGHNORMEFLGDSIMQIVATEYFIHPDHEHGLTLTLSSLYNNRT 1186
Db      953 AFLTLRTVGFNHLTLGHNORMEFLGDSIMQIVATEYFIHPDHEHGLTLTLSSLYNNRT 1012
Qy      1187 QAKVAEELGMOEYAITNDKTKRPGVGLRTKTLADLESFIALAYTDKOLEYVHTFMNVCFE 1246
Db      1013 QAKVAEELGMOEYAITNDKTKRPGVGLRTKTLADLESFIALAYTDKOLEYVHTFMNVCFE 1072
Qy      1247 PRLKPSIILNQMDNNDPSQLOQCCCLTRTSGKEPDILYKTLQTVGSHARTTVAVYFPG 1306
Db      1073 PRLKPSIILNQMDNNDPSQLOQCCCLTRTSGKEPDILYKTLQTVGSHARTTVAVYFPG 1132
Qy      1307 ERIGCGKSGISOAEKGAAMDALAEKNPFOMAHQKFKIGRKYROELKEMWREHERERP 1366
Db      1133 ERIGCGKSGISOAEKGAAMDALAEKNPFOMAHQKFKIGRKYROELKEMWREHERERP 1192
Qy      1367 DETEDIKK 1374

```

```

Db      1193 DETEDIKK 1200
RESULT 8
ABU08060
ID ABU08060 standard; protein; 1374 AA.
XX
AC ABU08060;
XX
DT 16-MAY-2003 (first entry)
XX
DE Human RNase III protein.
XX
KW Human; enzyme; RNase III; ribonuclease III; endoribonuclease;
KW pre-ribosomal RNA; pre-rRNA; small molecular weight nuclear RNA; snRNA;
KW small molecular weight nucleolar RNA; snorRNA; mRNA degradation;
KW antisense therapy; RNA interference; RNAi; gene therapy;
KW infectious agent; prophylaxis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region /label= Alpha_helix_1
FT Region /label= Alpha_helix_1
FT Region /label= Beta_sheet_1
FT Region /label= Beta_sheet_1
FT Region /label= Beta_sheet_2
FT Region /label= Beta_sheet_2
FT Region /label= Beta_sheet_3
FT Region /label= Beta_sheet_3
FT Region /label= Alpha_helix_2
PN US2002164601-A1.
XX
PD 07-NOV-2002.
XX
PF 06-JUL-2001; 2001US-00900425.
XX
PR 06-JUN-1996; 96US-00659440.
PR 06-JUN-1997; 97US-00870608.
PR 07-JUN-2000; 2000US-00479783.
XX
PA (WUHH/) WU H.
PA (CROO/) CROOKE S T.
PI Wu H, Crooke ST;
XX
DR WPI; 2003-328390/31.
XX
PT New human RNase polypeptide, useful for screening antisense
PT oligonucleotides for therapy of disorders associated with RNase III
PT expression or activity, or for evaluating the efficacy of an antisense
PT therapy.
XX
PS Claim 3; Fig 1; 17pp; English.
XX
XX The invention discloses an isolated human ribonuclease III (RNase III)
XX polypeptide, and the nucleic acid encoding it. RNase III is an
XX endoribonuclease that cleaves double stranded RNA. A1 RNase III species
XX contain an RNase III signature sequence. RNase III has been reported to
XX be involved in the processing of pre-ribosomal RNA (pre-rRNA), small
XX molecular weight nuclear RNAs (snRNAs) and small molecular weight
XX nucleolar RNAs (snorRNAs), as well as the degradation of some mRNA
XX species. Also disclosed is an antibody targeted to the human RNase III
XX polypeptide, an antisense compound 8 - 50 nucleobases in length, which is
XX targeted to the nucleic acid encoding human RNase III polypeptide, and
XX methods for inhibiting human RNase III expression, or activity, in a cell
XX or tissue, identifying agents that increase or decrease the activity or
XX levels of the human RNase III polypeptide in a host cell, screening
XX oligonucleotides to identify effective antisense oligonucleotides for
XX inhibition of expression of a selected target protein, prognosticating
XX efficacy of antisense therapy of a selected disease, eliciting cleavage

```

3
of a selected cellular RNA target and promoting RNA interference (RNAi)
in a cell. The RNase III polypeptide, the polynucleotide encoding it and
the antisense oligonucleotides, are useful for gene therapy (e.g. for
treating a disease or disorder associated with RNase III expression or
activity, or associated with an infectious agent), prophylaxis or as
research reagents. The sequence presented is the human RNase III protein
XX

SQ Sequence 1374 AA;

80.8%; Score 6062; DB 6; Length 1374;

Query Match
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MMQNTCHRMSPHPCGRCGRGCGHAGPAPSPFRPQNLRLHLPQCPVQYQYEPPSAPS 60
DB 274 MMQNTCHRMSPHPCGRCGRGCGHAGPAPSPFRPQNLRLHLPQCPVQYQYEPPSAPS 333
QY 61 TTFSSPAVNLPRPPDPVPPPPPPPSAQGLPCCPIPPPPNQMRRPPVPPPCFPMP 120
DB 334 TTFSSPAVNLPRPPDPVPPPPPPPSAQGLPCCPIPPPPNQMRRPPVPPPCFPMP 393
QY 121 PPPMPCPNPPVPGAPPCGCTFPFMMPPSPMHPPPPVPVMPQOVVQYPPGYSHNFPPP 180
DB 394 PPPMPCPNPPVPGAPPCGCTFPFMMPPSPMHPPPPVPVMPQOVVQYPPGYSHNFPPP 453
QY 181 SFNSFQNNSSFLPSANSSSPHFRLLPPYPLPKASERSRSPERLKHVDHRRHSHGR 240
DB 454 SFNSFQNNSSFLPSANSSSPHFRLLPPYPLPKASERSRSPERLKHVDHRRHSHGR 513
QY 241 GERHSLDRERGRSPRRQDSRYSDYDRGTRTPRRHSYRSRSPREERHRHRRNRS 300
DB 514 GERHSLDRERGRSPRRQDSRYSDYDRGTRTPRRHSYRSRSPREERHRHRRNRS 573
QY 301 PLEERSYKKEYSKGRSGSYGLSVPEPAGCTPELPGELIKNTSSWAPPLEIVNHRSPREK 360
DB 574 PLEERSYKKEYSKGRSGSYGLSVPEPAGCTPELPGELIKNTSSWAPPLEIVNHRSPREK 633
QY 361 KARWEEXKRWSDNOSGSKDNKNTSISKEPEETMPDKNEEBEELKPVWIRCTHSNN 420
DB 634 KARWEEXKRWSDNOSGSKDNKNTSISKEPEETMPDKNEEBEELKPVWIRCTHSNN 693
QY 421 YSSDPMDVGVSTVGTGRRLDLVDFEELGSRQEKAKARPPEPKTKLDELESS 480
DB 694 YSSDPMDVGVSTVGTGRRLDLVDFEELGSRQEKAKARPPEPKTKLDELESS 753
QY 481 SSECESDSDSTCSSSDSEVFVIAEIKRKKAHPRLDELWYNDPGQNDGPLCKSA 540
DB 754 SSECESDSDSTCSSSDSEVFVIAEIKRKKAHPRLDELWYNDPGQNDGPLCKSA 813
QY 541 KARRTGIRHSIYGEBAIKPCRPMTNNAQRLPHYRITVSPPTNFLTDRPTVIEYDHEXI 600
DB 814 KARRTGIRHSIYGEBAIKPCRPMTNNAQRLPHYRITVSPPTNFLTDRPTVIEYDHEXI 873
QY 601 FEGFSWFAAPLPLNLPLCKVIRNIDYTHFTEEMBPENFCVKGLEFLFPRDLLEY 660
DB 874 FEGFSWFAAPLPLNLPLCKVIRNIDYTHFTEEMBPENFCVKGLEFLFPRDLLEY 933
QY 661 DNLIKGPLFEDSPPCPCPFHFMFRFVRLPDGKEVLSHQILLYLRSKALVPEEBIA 720
DB 934 DNLIKGPLFEDSPPCPCPFHFMFRFVRLPDGKEVLSHQILLYLRSKALVPEEBIA 993
QY 721 NMLQWELMOKYAECKGMIVTNPCTKSSVRIIDLREQNPDVITPIIVHFGIRPA 780
DB 994 NMLQWELMOKYAECKGMIVTNPCTKSSVRIIDLREQNPDVITPIIVHFGIRPA 1053
QY 781 QLSYAGDPVOYKLMKSYVLRHLNLSPKYKQDQKLAQREBALQKIKQKTMREEVV 840
DB 1054 QLSYAGDPVOYKLMKSYVLRHLNLSPKYKQDQKLAQREBALQKIKQKTMREEVV 1113
QY 841 ELSSQGFMTGTGRSDVCOHAMMLPVLTHHRIHYQCMLHDLKIGYFODRCILLQMTMP 900
DB 1114 ELSSQGFMTGTGRSDVCOHAMMLPVLTHHRIHYQCMLHDLKIGYFODRCILLQMTMP 1173
```

```
QY 901 SHHLNFGMNPDAHNSLNSCGIROPKYGRKVVHNMKKGKINTLINTMSRLGODPPTS 960
DB 1174 SHHLNFGMNPDAHNSLNSCGIROPKYGRKVVHNMKKGKINTLINTMSRLGODPPTS 1233
QY 961 RINHERLEFLGDAVVEFLTSVHLYLFPSPLEGGATATRTAVONQHLAMLAKLELDP 1020
DB 1234 RINHERLEFLGDAVVEFLTSVHLYLFPSPLEGGATATRTAVONQHLAMLAKLELDP 1293
QY 1021 FMLYAHGPDLCRESDLRHMANMCFPALIGAVYLEGSLERAKOLFGLLFNDPDLREWLN 1080
DB 1294 FMLYAHGPDLCRESDLRHMANMCFPALIGAVYLEGSLERAKOLFGLLFNDPDLREWLN 1353
QY 1081 YPLHPLQLOEPTDRLIETS 1101
DB 1354 YPLHPLQLOEPTDRLIETS 1374
```

RESULT 9

ID AAB92635 standard; protein; 769 AA.

AC AAB92635;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:10949.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-

XX length cDNAs defined in the specification, and for the detection and/or

XX diagnosis of the abnormality of the proteins encoded by the full-length

XX cDNAs.

XX Claim 8; SEQ ID NO 10949; 2537bp + Sequence listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-

XX length cDNAs defined in the specification. Where a primer set comprises:

XX (a) an oligo-dT primer and an oligonucleotide complementary to the

XX complementary strand of a polynucleotide which comprises one of the 5602

XX complementary sequences defined in the specification, where the

XX complementary sequence comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in the

XX specification. The primer sets can be used in antisense therapy and in

XX gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX Sequence 769 AA;

Query Match 54.3%; Score 4071; DB 4; Length 769;
Best Local Similarity 99.5%; Pred. No. 4,2e-280;
Matches 765; Conservative 0; Mismatch 4; Indels 0; Gaps 0;

QY 606 MFAHAPLTVNIPKCVIRFNIDYTHIFIEEMMPENFCVKGLESLFLFRDILELYDMNLK 665
DB 1 MFAHAPLTVNIPKCVIRFNIDYTHIFIEEMMPENFCVKGLESLFLFRDILELYDMNLK 60
QY 666 GRLFEBSPPCCPRFHFMPFRFRLPDGKEVLSHQLLTLRLCSKALVBEERIANNLQW 725
DB 61 GRLFEBSPPCCPRFHFMPFRFRLPDGKEVLSHQLLTLRLCSKALVBEERIANNLQW 120
QY 726 EELMOKVAEECGMTVTNPGTSSVRIDQLDRBOPNPVITFPIIVHGIIPAOLSYA 785
DB 121 EELMOKVAEECGMTVTNPGTSSVRIDQLDRBOPNPVITFPIIVHGIIPAOLSYA 180
QY 786 GDFQYOKLWKSYYKLHLLANSPKVKQTDKQKLAQREALQKTRKNTMRREYTVELSSQ 845
DB 181 GDFQYOKLWKSYYKLHLLANSPKVKQTDKQKLAQREALQKTRKNTMRREYTVELSSQ 240
QY 846 GFAPKTSIRSDVCOHMMPLVLTHTIRYHOCMLHDLKLTGYTFODRCILLQAMTHPSHHLN 905
DB 241 GFAPKTSIRSDVCOHMMPLVLTHTIRYHOCMLHDLKLTGYTFODRCILLQAMTHPSHHLN 300
QY 906 FGNNPRAHRSLSNCGIRORPKYGDRAKVNHNMRKKGINTLIMSRUGODDPTPSRIHNH 965
DB 301 FGNNPRAHRSLSNCGIRORPKYGDRAKVNHNMRKKGINTLIMSRUGODDPTPSRIHNH 360
QY 966 ERLEFLGDAAVEEFLTSVHLYYLFPSSLEEGSLATYRTAIVONQHILAMLAKELEDPFMYA 1025
DB 361 ERLEFLGDAAVEEFLTSVHLYYLFPSSLEEGSLATYRTAIVONQHILAMLAKELEDPFMYA 420
QY 1026 HGPDLGRESDDLHMANCFEALIGAVYLEGSLSEBAKQLFGRLLPNDPLREVMILNPLHP 1085
DB 421 HGPDLGRESDDLHMANCFEALIGAVYLEGSLSEBAKQLFGRLLPNDPLREVMILNPLHP 480
QY 1086 LQIQENTDROQLIETSPVLOKLTPEFEALIGVITTHYRLARATTLRVGNHNLTLGNOR 1145
DB 481 LQIQENTDROQLIETSPVLOKLTPEFEALIGVITTHYRLARATTLRVGNHNLTLGNOR 540
QY 1146 MEFLGDSIMQVATEVLELFIHPDHHECHLTLLRSSLVNNRTQAKVAEELMOEYATINDK 1205
DB 541 MEFLGDSIMQVATEVLELFIHPDHHECHLTLLRSSLVNNRTQAKVAEELMOEYATINDK 600
QY 1206 TKRPVGLRTKTLADLESFIALYTDKLELYVTFEMNVCFPRLKEFILNQDNDPKSQ 1265
DB 601 TKRPVGLRTKTLADLESFIALYTDKLELYVTFEMNVCFPRLKEFILNQDNDPKSQ 660
QY 1266 QQCCLTLRTBEGKRPDPLKYTLQTVGSHARTTVAVYFKGERIIGCGKGSIOQAEMGAA 1325
DB 661 QQCCLTLRTBEGKRPDPLKYTLQTVGSHARTTVAVYFKGERIIGCGKGSIOQAEMGAA 720
QY 1326 MDALFKYNFQMAHOKRFGIKRYROELKEMRWEREHOEREPDETBDIKK 1374
DB 721 MDALFKYNFQMAHOKRFGIKRYROELKEMRWEREHOEREPDETBDIKK 769

RESULT 10
AAB58539
ID AAB58539 standard; protein; 1327 AA.
AC AAB58539;
XX
DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 2409.
XX KM Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN MO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001MO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL02642.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.

PS Disclosure; SEQ ID NO 2409, 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AAB57737-
CC AAB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1327 AA;

Query Match 39.5%; Score 2963; DB 4; Length 1327;
Best Local Similarity 45.6%; Pred. No. 4,7e-201;
Matches 621; Conservative 213; Mismatch 327; Indels 200; Gaps 32;

QY 90 QGFLPPCPPIRPPRNHQMHPFVPPCFPPMPRPPCPN-NRPVGAAP-----P 137
DB 3 QGFLPPCPPIRPPRNHQMHPFVPPCFPPMPRPPCPN-NRPVGAAP-----P 137
QY 138 GQGTFFPMMRPPRPPPPPPVPPQVNVYQPRGY-----SHNFP 179
DB 47 SSKSLDYVPEPAPYASS---VPSYDYPQAPAYEGEYAVNEQAOKTGGQSHYQY 103
QY 180 PSFNSPQNNPSPFLPANSSSPHNLPPYPL---PKAPSERRSPERLKHVDHRRHD 235
DB 104 PA-----SGSSTLYS-----YKPRDLYPSSNYPSPRQ-----RYTS 139
QY 236 HSHGGERHRSLLDRRGRSPDRRQDSRYRSDYDRGRTSPSRHSYRSRERERHRR 295
DB 140 NSSGQYHNP-----GYSSGR-----RYEQHNDQ-----EHRQLODSRYAHEPRGHY 183
QY 296 DNRSPSLERSYKKEYKRSRSGSLGVPEPACGTBELPQELIKNTDMSAPRLEYNNHS 355
DB 184 AHRQAQSGOHY---YGAARN-----QVSDDSYRSGHNEREN 219
QY 356 PSREKGRAC--RWEERKOR-----WSDNOSGKDKNTYSIKEPEEPTMPDNBESEELK 409
DB 220 ETLEKTRAKRYETEDRLRLQWCSNFC-----EKEDYVKKGNALSEADAV 267
QY 410 PWNIRCTSHSENNYSSDPMQOVGDTYVGTGRLLDYDKPEEELGSGOEKAAKAPRWEPP 469


```

Db      268 ESWVRSSPAELYERTK-----SENEVRGARLQKCLTLPDELLQRAKVRKLPVYVP 323
Qy      470 KTKL-----DEDLSSSESECEDESDTSCSSSDSEVPDV-----IAEIKKKAH 514
Db      324 PKRARRVCVKHKHSEACSSSSSSDDSDDEDA-----FKIQQCCMEELSKRVQH 373
Qy      515 PDRLHBLWYNDGQWNDGFLCKCSAKARTGIRHSIYPEGEALIKRCPMTNNAGLFLFY 574
Db      374 PQRVHADLWHDNDGEMNDGFLCRCSAKSRIRIGRIHIIYPEGTGYKLCDPSSNAGLFLFY 433
Qy      575 RITVSPPTNPLTRPVIEYEDDEHYIPEGSFMAHAPLTNIPLCVIRFNIDYTHIEE 634
Db      434 RISISPTNPLTPTPIIKDHEHPLPEGSLSLHRLSLPVCXIRIRNIETIYESEE 493
Qy      635 MPEPENCYKGLLESLPLFRDILELYDWMKGPL-----FEEDSPCCPRHFMRFPVRLP 690
Db      494 KQPEPNTIHELDIFPKYLFHELELVDPNLMPLPGSNVBS---CPAFHFFPRFVRLP 550
Qy      691 DGGKEVLSHQIILLYLRGSKALVPEBELANMLQMELEKQYAECKGMIVNPGTKPS 750
Db      551 DNGKEVLAWEVLRYYLLDNSAQLVERQQLHLNQISOEQWNTVDTIKGLVYKPGYKEC 610
Qy      751 SVRIDDLDEQFPNDV-----ITPPIYHFGIRPAQLSYAGDPPOYOKLMSVYKLR 801
Db      611 SLRVQDLDRN--NSDLPECVDRGTGSHRAIVHFGICHPOLSYAGDPYOKAMREYVYKTR 668
Qy      802 HLANSRKVQTDQKQLAQREBALQKIROKNTMRREVVELSSQGFMTGIRSDVCQHAM 861
Db      669 HLMANNSKSPFQDKRLSEKEQRLQEMRQTGMRKRNITVAISSEGVYRGIWGDVQHAM 728
Qy      862 MLPVLTNIRYHOCMLMDKLIQYTFQDRCLDLQAMTHSHNLNFGMNDPHANSLNSNG 921
Db      729 LIPVLGHRFHKSLDLLEESIGYRFKNRYLQALATHSYKKNYVNTNPHANSLNTNG 788
Qy      922 IQPQKGDGRVHNMHRKKGINTLINIMSRIGQDDPTPSRINNESEFLGDAVVEFLTS 981
Db      789 IQPQKGDGRKINMTRKGINLTVSISMRFGENHETVSNITNNESEFLGDAVVEFLTS 848
Qy      982 VHLVYLPSPLEEGSLATYRTAIVQNHMLAKKELDPRMLYAHGPDLCRESDLRHAMA 1041
Db      849 IHLFMRPELEEGSLATYRTAIVQNHMLAKKELDPRMLYAHGSDLCHELELRHAMA 908
Qy      1042 NCFEALIGAVYLEGSLEBAKOLFGRLLF--NDPRLREVWMLNPLHPIQLOEPNTDROLIET 1100
Db      909 NCFEALMGALLLDGGIKVADDEVFTDALFRDEKLLSIWKLPEHPQOEPRLDGDSCLDS 968
Qy      1101 SPVLQKLTPEBEAIGVIFTHVRLARAFTLRTVGFNNHLLTGHQRMPELGDSTIMOLVATE 1160
Db      969 YRVLKELTPEEDSIGIKFKHIRLARAFTDRSIGFTHLTLGNSQRLSEFLGDTVLQICS 1028
Qy      1161 YLFHFRDHHGHLLTLRSLVNNRTQAKVAEELGMOEYAI--TNDKTKRPVGLRTKTLAD 1219
Db      1029 YLVRHREHNEHGLSLRSLVNNRTQAVVCCDLGMRPKAVVANPR-----ADLKTDRAD 1084
Qy      1220 LLESFIALYTDKOLEVYNTFMVNCFFRPLKEITLQDMNDPKSQLOCCCLTIRT--EGKE 1278
Db      1085 LLEAFIGALYVDKGLLYCEQFCFVCLFPRLOLFIMQDMNDPKSKLOCCCLTLRTMDGSE 1144
Qy      1279 PDIPLVKTQVPSHARTYTVAVVFKSGEIGGKSPSTQQAAMGAAMALEKYN--PFO 1336
Db      1145 PDIPLVKKVVASPNTYTRKYAVAVVFRSKLATSSSSSTQQAAMNAKQALENSRDLFPQ 1204
Qy      1337 MAHQKRFGRKYRQEL---KEMWEREHOERBEDETDIKK 1374
Db      1205 LDHQKRVIAKSIKKQTGNELDNDSRQHDE-----EKIKR 1239

```

```

RESULT 11
ABM04841
ID      ABM04841 standard; protein; 541 AA.
XX      ABM04841;
AC      ABM04841;
XX

```

```

DT      22-SEP-2003 (first entry)
XX
DE      Human putative ribonuclease III.
XX
KW      spinal cord; neuropathic pain; central sensitisation pain; pain;
XX      analgesic; gene therapy.
XX
OS      Homo sapiens.
XX
PN      EP1284298-A2.
XX
PD      19-FEB-2003.
XX
PF      26-JUL-2002; 2002EP-00255229.
XX
PR      27-JUL-2001; 2001GB-00018354.
PR      07-FEB-2002; 2002GB-00002883.
XX
PA      (WARN ) WARNER LAMBERT CO.
XX
PI      Brooksbank RA, Dixon AK, Lee K, Plincock RD;
XX
DR      MPI; 2003-543489/52.
DR      N-PSDB; ACF25385.
XX
PT      Use of an isolated gene sequence in the screening of compounds for
PT      diagnosing or treating pain.
PS      Claim 1; Page 183-184; 188pp; English.
XX
CC      The invention relates to a novel isolated gene sequence that is
CC      downregulated in the spinal cord of a mammal in response to mechanically
CC      distinct first and second models of neuropathic or central sensitisation
CC      pain, useful in the screening of compounds for diagnosing or treating
CC      pain. A protein encoded by a gene of the invention has analgesic
CC      activity. A polynucleotide of the invention may have a use in gene
CC      therapy. The gene sequence is useful for preparing a composition for
CC      diagnosing or treating pain. The present sequence represents a protein
CC      encoded by a gene of the invention
XX
SQ      Sequence 541 AA;
XX
Query Match      38.0%; Score 2847; DB 6; Length 541;
Best Local Similarity 99.3%; Pred. No. 2.5e-193;
Matches 537; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy      834 MREVTVELSSQGFMTGIRSDVCOHAMMLPVLTNHIRYHOCMLMDKLIQYTFQDRCLL 893
Db      1 MREVTVELSSQGFMTGIRSDVCOHAMMLPVLTNHIRYHOCMLMDKLIQYTFQDRCLL 60
Qy      894 QLAMTHPSHNLNFGMNDPHARNSLNSCGIRQPKYGDGRKVHNMHRKKGINTLINIMSRIG 953
Db      61 QLAMTHPSHNLNFGMNDPHARNSLNSCGIRQPKYGDGRKVHNMHRKKGINTLINIMSRIG 120
Qy      954 QDDPFPISRINHEBERFLGDVAVEFLTSVHLVYLPSPLEEGSLATYRTAIVQNHMLAMA 1013
Db      121 QDDPFPISRINHEBERFLGDVAVEFLTSVHLVYLPSPLEEGSLATYRTAIVQNHMLAMA 180
Qy      1014 KGLLEDPFVLVYHGPDLCEESDLRHAMANCFEALGAVYLEGSLEBAKOLFGRLLFNDDP 1073
Db      181 KGLLEDPFVLVYHGPDLCEESDLRHAMANCFEALGAVYLEGSLEBAKOLFGRLLFNDDP 240
Qy      1074 LREVMNLNPLHPILOQEPNTDROLIETSPVLQKLTPEBEAIGVIFTHVRLARAFTLRTV 1133
Db      241 LREVMNLNPLHPILOQEPNTDROLIETSPVLQKLTPEBEAIGVIFTHVRLARAFTLRTV 300
Qy      1134 GFNNHLLTGHQRMPELGDSTIMOLVATEYLFIFPDHHEGHLLTLRSLVNNRTQAKVAAB 1193
Db      301 GFNNHLLTGHQRMPELGDSTIMOLVATEYLFIFPDHHEGHLLTLRSLVNNRTQAKVAAB 360
Qy      1194 LGMQEYAITNDKTKRPVGLRTKTLADLLESFIALYTDKOLEVYNTFMVNCFFRPLKEFI 1253
Db      361 LGMQEYAITNDKTKRPVGLRTKTLADLLESFIALYTDKOLEVYNTFMVNCFFRPLKEFI 420

```

OY		125	LNDQWNPDKSLOOCCLLFRGEKPEPDIPLKXTLTQTVPSHARTYTAVVRKGRIQCGK	13313
Dd		421	LNDQWNPDKSLOOCCLLFRGEKPEPDIPLKXTLTQTVPSHARTYTAVVRKGRIQCGK	480
OY		1314	GPSIQQAEMGAAAPDALEKYKNFPQMAHQRFIQRKTKRQLKEMRWEREHOEREPDETEDIK	13733
Dd		481	GPSIQQAEMGAAAPDALEKYKNFPQMAHQRFIERKTKRQLKEMRWEREHOEREPDETEDIK	540
OY		1374	K 1374	
Dd		541	K 541	
RESULT 12				
ID	AEA21037			
ID	AEA21037	standard; protein; 486 AA.		
AC	AEA21037;			
XX				
DT	11-AUG-2005	(first entry)		
DE				
XX				
XX				
KM		Novel human polypeptide SEQ ID NO 1731.		
KM		vulnerary; CNS-gen.; gene therapy; diagnostic; forensic; mapping;		
KM		DNA purification; protein purification; osteoarthritis; antiarthritic;		
KM		osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;		
KM		periodontal disease; antiinflammatory; mouth disease; burns; injury;		
KM		peripheral neuropathy; Alzheimer's disease; neuroprotective; nootropic;		
KM		dementia; parkinson's disease; antiparkinsonian; neurological disease;		
KM		cerebrovascular ischemia; cerebroprotective; vasotrophic;		
KM		cardiovascular disease; autoimmune disease; immunosuppressive;		
KM		immune disorder; viral infection; vitruicide; infection; cancer;		
KM		cyostatic; neoplasm.		
XX				
OS	Homo sapiens.			
XX				
PN	WO2005049806-A2.			
PD	02-JUN-2005.			
PF	11-MAR-2004; 2004WO-US007412.			
PR	14-MAR-2003; 2003US-00389559.			
PA	(NUVE-) NUVELO INC.			
PI	Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y,			
PI	Gao M, Xue A, Auand V, Zhao Q, Wang D, Goodrich R, Chen R,			
PI	Wehrman T, Meng G, Boyle B,			
DR	WPI; 2005-417730/42.			
PT				
PT	New polynucleotide encoding a polypeptide with biological activity,			
PT	useful for treating a disease or disorder, e.g. osteoarthritis, burns,			
PT	CNS and peripheral disease, stroke, autoimmune disorders, viral			
PT	infection, or cancer.			
PS				
XX	Example 3; SEQ ID NO 1731; 500bp; English.			
CC				
CC	The invention describes a new isolated polynucleotide (1) encoding a			
CC	polypeptide with biological activity comprising: a nucleotide sequence of			
CC	SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes			
CC	to the sequence of (1) under stringent hybridization conditions; or a			
CC	nucleotide sequence having greater than 9% sequence identity with the			
CC	sequence of (1). Also described are: a(n) (expression)vector comprising			
CC	(1); a host cell genetically engineered to comprise (1) operatively,			
CC	associated with a regulatory sequence that modulates expression of the			
CC	polynucleotide in the host cell; an isolated polypeptide comprising a			
CC	*sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide			
CC	18; a polypeptide encoded by (1); or a polypeptide encoded by a			
CC	polynucleotide hybridizing under stringent conditions with any one of SEQ			
CC	ID NOS: 1-567; a composition comprising the polypeptide of (3) and a			
CC	ID NOS: 1-567; a composition comprising the polypeptide of (3) and a			

CC carrier, an antibody directed against the polypeptide of (3); a method
CC for detecting (1) in a sample; a method for detecting the polypeptide of
CC (3) in a sample; a method for identifying a compound that binds to the
CC polypeptide of (3); a method for producing the polypeptide of (3); and a
CC collection of polynucleotides, where the collection comprising of at
CC least one of SEQ ID NOS: 1-567. (1) is a polynucleotide comprising any of
CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological
CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:
CC 568-1134. All sequences are fully defined in the specification. The
CC sequences and methods are useful in diagnostics, forensic, and gene
CC mapping, in identifying of mutations responsible for genetic disorders or
CC other traits, in assessing biodiversity, and for producing many other
CC types of data and products dependent on DNA and amino acid sequences. The
CC composition and method are useful for treating a disease or disorder,
CC e.g., osteoporosis, osteoarthritis, periodontal disease, burns, CNS and
CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke, and
CC autoimmune disorder, viral infection, or cancer. This is the amino acid
CC sequence of a novel polypeptide of the invention.
XX
SQ Sequence 486 AA;

	Query Match	33.6%	Score 2519;	DB 9;	Length 486;
	Best Local Similarity	99.2%	Pred. No.4.4e-170;		
	Matches 477;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0
Qy	QIATHTPSHHLPFGNPNPDHANSLNSCGIRPKYGDGRKYVHHMNRKKGINTLINIMSLRG	894			953
Db	QIATHTPSHHLPFGMNPDPHANSLNSCGIRPKYGDGRKYVHHMNRKKGINTLINIMSLRG	6			65
Qy	QDDPTPSRININERLEFLGDAVVEFLTSVHLYYLFPSSLSEGGATYRTAIYVONHMLMLA	954			1013
Db	QDDPTPSRININERLEFLGDAVVEFLTSVHLYYLFPSSLSEGGATYRTAIYVONHMLMLA	66			125
Qy	KKLEIDRPMLYAHGPDLCRESDLNHMANCFEALIGAVYLEGSLSEBAKOLFGRLIFNDPD	1014			1073
Db	KKLEIDRPMLYAHGPDLCRESDLNHMANCFEALIGAVYLEGSLSEBAKOLFGRLIFNDPD	126			185
Qy	LREVLNLYPLHPLQLOEPTNDROLIETSPVLOKLTPEFEAIGVFTYHRLARAFLLRTV	1074			1133
Db	LREVLNLYPLHPLQLOEPTNDROLIETSPVLOKLTPEFEAIGVFTYHRLARAFLLRTV	186			245
Qy	GFNHLTLGHNRMEFLGDSIMQVATEYLFPIHPDHHGHLLTLRSSLVNNRTOAKVAEE	1134			1193
Db	GFNHLTLGHNRMEFLGDSIMQVATEYLFPIHPDHHGHLLTLRSSLVNNRTOAKVAEE	246			305
Qy	LGMOEYAITNDKTRPVGLRTKTLADLLESLIAYLDKOLEYVHTFMNVCFFPRLKEFI	1194			1255
Db	LGMOEYAITNDKTRPVGLRTKTLADLLESLIAYLDKOLEYVHTFMNVCFFPRLKEFI	306			365
Qy	LNDQMDPKSLOQCCLTLRTEGKEPDIPLKYTLQTVGSPHARTYTAAVYFKGERIGGK	1254			1313
Db	LNDQMDPKSLOQCCLTLRTEGKEPDIPLKYTLQTVGSPHARTYTAAVYFKGERIGGK	366			425
Qy	GPSIQOAEEMGAMDALEKYNPPQAHQGRFLGRKYRQELKEMRWERHQBREPDETEDIK	1314			1373
Db	GPSIQOAEEMGAMDALEKYNPPQAHQGRFLGRKYRQELKEMRWERHQBREPDETEDIK	426			485
Qy	1374 K 1374				
Db	486 K 486				
RESULT 13					
ADQ96710					
ID	ADQ96710 standard; protein, 466 AA.				
XX	ADQ96710;				
AC					
XX					
DT	23-SEP-2004 (first entry)				
XX					
DE	Human ribonuclease III (RNase III)-like domain.				
XX					
KX	Human; ribonuclease III; RNase III; research purpose; biological purpose;				

KW clinical purpose; cellular interaction.
XX
OS Homo sapiens.
XX
PN US2004126867-A1.
XX
PD 01-JUL-2004.
XX
PF 09-FEB-2004; 2004US-00774974.
XX
PR 06-JUL-2001; 2001US-00900425.
XX
PA (CROO/) CROOKE S T.
XX
PA (WUHH/) WU H.
XX
PI Crooke ST, Wu H;
XX
DR WPI, 2004-516913/49.
XX
PT New isolated nucleic acid molecule encoding human RNase III, useful for
PT research, biological, or clinical purposes, e.g. defining the roles of
PT RNase III and the interaction of human RNase III and cellular RNA.
XX
PS Example 9, SEQ ID NO 37; 31pp; English.
XX
XX The invention relates to human ribonuclease III (RNase III) and its
CC corresponding nucleic acid sequence. The polynucleotide sequence of the
CC invention is useful for research, biological and clinical purposes. It is
CC useful in defining the roles of RNase III and the interaction of human
CC RNase III and cellular RNA. The present sequence is human RNase III-like
CC domain protein.
XX
SQ Sequence 466 AA;

Query Match 32.8%; Score 2463; DB 8; Length 466;
Best Local Similarity 100.0%; Pred. No. 4e-166;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 909 NPDHANSNSGCGIRPKYKGRVHMHMRKKGINTLINIMSRIGODDTPPSRINNEHL 968
DB 1 NPDHANSNSGCGIRPKYKGRVHMHMRKKGINTLINIMSRIGODDTPPSRINNEHL 60
QY 969 EFLGDAVVEFLTSVHLVYLFPSSIEEGGLATYRTAIVQNHMLAKKLEDPFMYAHGP 1028
DB 61 EFLGDAVVEFLTSVHLVYLFPSSIEEGGLATYRTAIVQNHMLAKKLEDPFMYAHGP 120
QY 1029 DICRESDLAHAMANCEEALIGAVYLEGSLBEAKQLFGRLFPNDPDLREYIWPILHPLQL 1088
DB 121 DICRESDLAHAMANCEEALIGAVYLEGSLBEAKQLFGRLFPNDPDLREYIWPILHPLQL 180
QY 1089 QEPNTDROLIETSPVLOKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHNORMEF 1148
DB 181 QEPNTDROLIETSPVLOKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHNORMEF 240
QY 1149 LQDSIMQVATEYLFTHFPDHEGHLLTLRSSLVNNRTQAKVAEBIGMOEYAITNDKTKR 1208
DB 241 LQDSIMQVATEYLFTHFPDHEGHLLTLRSSLVNNRTQAKVAEBIGMOEYAITNDKTKR 300
QY 1209 PVGLRKTTLADLIESFIALYTDKDELYVHTFMNVCFEPRLKEFTLINQDMNPKSLOQC 1268
DB 301 PVGLRKTTLADLIESFIALYTDKDELYVHTFMNVCFEPRLKEFTLINQDMNPKSLOQC 360
QY 1269 CULTRTEGKEPDPPLKTLQTVGSHARTYTVAVVFKGERIGCGKPSIQAEEMGAAMA 1328
DB 361 CULTRTEGKEPDPPLKTLQTVGSHARTYTVAVVFKGERIGCGKPSIQAEEMGAAMA 420
QY 1329 LEKYNPQMAHQKRFGRKYRQELKEMRERHQEREPEDEDIKK 1374
DB 421 LEKYNPQMAHQKRFGRKYRQELKEMRERHQEREPEDEDIKK 466

RESULT 14
AAU20587

ID AAU20587 standard; protein; 378 AA.
XX
AC AAU20587;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human secreted protein, Seq ID No 579.
XX
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.
XX
OS Homo sapiens.
XX
PN WO200155326-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001347.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198122P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0226924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
PI
XX
XX WPI; 2001-451931/48.
DR N-PSDB; AAS33296.
XX
XX
PT New nucleic acids and polypeptides, useful for diagnosing, preventing or
PT treating medical conditions.
XX
XX
PS Claim 11; SEQ ID NO 579; 753bp; English.
XX
XX
CC The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 26.1%; Score 1961; DB 4; Length 378;
Best Local Similarity 98.9%; Pred. No. 1,4e-130;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 997 ATYRTAIVONOHILAKKLEIDPFMLYAHGPDLCRESDLRIAMANGFALIGAVYLEG 1056
DB 1 ATYRTAIVONOHILAKKLEIDRFMLYAHGPDLCRESDLRIAMANGFALIGAVYLEG 60
QY 1057 LEEAKOLRGRLLEPNPDLEWLVNPLHLOLOEPTDQOLLETSVYLOKTEFEPAICV 1116
DB 61 LEEAKOLFRLLEPNPDLEWLVNPLHLOLOEPTDQOLLETSVYLOKTEFEPAICV 120
QY 1117 IFTHYRLARAFLLRTVGFNHLLTIGNORMEFLGDSIMOLVATXYLFIFPDHHEGHTL 1176
DB 121 IFTHYRLARAFLLRTVGFNHLLTIGNORMEFLGDSIMOLVATXYLFIFPDHHEGHTL 180
QY 1177 LRSLVNNRTQAKVAEELGMOEYAITNDTKRPVGLRTYTLADLLESFALYTDKLEY 1236

Db 161 LRSILNNKTOAKVABELGMSYAITNDKTRPVALRTTLADLESFIAALYIDKOLEY 240
QY 1237 VHTFNNVCFPRILKEPIILNQDNNDPKSQIQCCCLTLRTGKEPDIPLYKTLQTVGSPSHAR 1296
Db 241 VHTFNNVCFPRILKEPIILNQDNNDPKSQIQCCCLTLRTGKEPDIPLYKTLQTVGSPSHAR 300
QY 1297 TTYTVAVYFKGERIGCGKGPSIQAEWGAAMDALKEKNFPMQAHQKRFIRKTRQSLKEMR 1356
Db 301 TTYTVAVYFKGERIGCGKGPSIQAEWGAAMDALKEKNFPMQAHQKRFIRKTRQSLKEMR 360
QY 1357 WEREHOERPEDETDIKK 1374
Db 361 WEREHOERPEDETDIKK 378

RESULT 15
AAU21744
ID AAU21744 standard, protein, 378 AA.
XX
AC AAU21744;
XX
DT 04-DEC-2001 (first entry)
XX
DE Novel human neoplastic disease associated polypeptide #177.
XX
KW Human; neoplastic disease associated polypeptide; cancer;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasoactive.
OS Homo sapiens.
XX
XX WO200155163-A1.
PN
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001358.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227099P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: February 3, 2006, 19:36:34 ; Search time 55 Seconds
(without alignments)
2403.670 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500

Sequence: 1 NM0GNTCHMSFHPGRGCRP.....MRWERHOREPDETDIK 1374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	890	11.9	412	2 T21419	hypothetical prote
2	558	7.4	682	2 T21420	hypothetical prote
3	316	4.2	551	2 S57447	HPBRI-7 protein -
4	315	4.2	249	2 B69693	ribonuclease III (
5	307.5	4.1	229	2 AE1300	ribonuclease III h
6	300	4.0	229	2 AE1672	ribonuclease III h
7	293	3.9	1560	2 T42727	proliferation pote
8	285.5	3.8	225	2 B82073	ribonuclease III v
9	282	3.8	263	2 A83961	ribonuclease III r
10	281.5	3.8	232	2 P98012	ribonuclease III (
11	279.5	3.7	232	2 H95144	ribonuclease III (
12	276	3.7	948	2 A57640	retinoblastoma bin
13	275	3.7	760	2 T06291	extensin homolog T
14	274.5	3.7	1531	2 T48946	hypothetical prote
15	273.5	3.6	1006	2 G86292	hypothetical prote
16	273	3.6	891	2 G84693	probable prolina-r
17	271.5	3.6	620	2 G06733	hydroxyproline-ric
18	270	3.6	226	2 H84959	ribonuclease III (
19	268.5	3.6	231	2 B86725	ribonuclease III (
20	267	3.6	224	2 A81260	ribonuclease III (
21	267	3.6	243	2 D89896	RNase III (importe
22	267	3.6	998	2 T30930	hypothetical prote
23	265.5	3.5	1002	2 T43236	carboxypeptidase C
24	262	3.5	226	1 NRE63	ribonuclease III (
25	262	3.5	226	2 P85902	RNase III, de RNA
26	262	3.5	226	2 A91058	extensin precursor
27	259	3.5	322	2 S25299	hypothetical prote
28	258	3.4	231	2 S76204	hypothetical prote
29	257.5	3.4	272	2 T35656	ribonuclease III -

30	256	3.4	246	2 H70187	ribonuclease III (
31	255.5	3.4	1638	2 A42091	transcription acti
32	255	3.4	226	2 AC0829	ribonuclease III (
33	254	3.4	1611	2 T38236	hypothetical prote
34	253.5	3.4	368	2 C29356	hydroxyproline-ric
35	253.5	3.4	1048	2 T31425	C-terminal domain-
36	252	3.4	226	2 AF0331	ribonuclease III (
37	251.5	3.4	240	2 F64602	ribonuclease III (
38	251.5	3.4	1201	2 G86441	unknown protein (I
39	250.5	3.3	727	2 C84534	hypothetical prote
40	249	3.3	225	2 T09964	extensin CYC15 pre
41	249	3.3	907	2 E96836	hypothetical prote
42	249	3.3	1877	2 T21861	hypothetical prote
43	245	3.3	230	2 G97115	dsRNA-specific rib
44	244.5	3.3	489	2 T11622	extensin class 1 p
45	243.5	3.2	708	2 D96711	hypothetical prote

ALIGNMENTS

RESULT 1
T21419
hypothetical protein F26B4.13 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T21419
R/Highlighting, J.
Submitted to the EMBL Data Library, October 1996
A/Reference number: Z19419
A/Accession: T21419
A/Status: Preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-412 <NLS>
A/Cross-references: UNIPARC:UPI000017B936; EMBL:Z81070; PIRN:CAB03005.1; GSPDB:GN00019; C
A/Experimental source: clone F26B4
C/Genetics:
A/Gene: CSP: F26B4.13
A/Map position: 1
A/Introns: 359/2; 400/3

Query Match	11.9%	Score 890;	DB 2;	Length 412;
Best Local Similarity	43.1%;	Pred. No. 6.6e-40;		
Matches 181;	Conservative 85;	Mismatches 136;	Indels 18;	Gap 7;
QY	944	TLINIMSRIGQDDPTSRINHNRELEFLGDVAVVEFLTSVHLYTLFSLBEGGLATRTAI	1003	
DB	2	SLFNIMKGTSGGP---ILHNERLEYLDVAVVELVSHLYFMTLHFEGLATYRTAL	57	
QY	1004	VONQHLAMLAKKLELDPFMYAHGPDLGRESDLRHMANCFEALIGAVVLEGLSEBAKOL	1063	
DB	58	VONRNLAATLAKNCRIDEMLOISHGADLVINAEKALIANFEVMAIYLDGLACDVI	117	
QY	1064	FGRLIF-NDPDLREVLNPLPLQLQBPNTDROLIETSPVLOKTEFEBAIGVIFTHVR	1122	
DB	118	FSKAVVGHQPVLEKKKDHINEHELKREDDPGDRDLSFIPPTLSTFALBERRIGQNNIR	177	
QY	1123	LARARLTTLTVGNHNLTLGHNQMEFLGDSIMQVATEVLFIFHPDHGHCHLTLLASSLV	1182	
DB	178	LAKAFTRRNIPNDLTKGHNOFLLEWLGDSVLOLYSDPLYRFPFHGHGMSLTLTSIV	237	
QY	1183	NMRTOAKVAEBELGMOEYATINDTKRPV---GLRTYLLDLLESFALAYLTDDLEAVNT	1239	
DB	238	SNQTOAVVCDLDGTFEFLV----KAPYTPPELKKDKADLVFAFGALYVDRGIIHCR	292	
QY	1240	FMNVCFPPRLKEBETLQDNWNPDSQLOQCCLTLR-TEGKEPDIPLYKTLQTVGSPHARTV	1298	
DB	293	FRIIVFCPLKPISEKKNDKASHLQWCCLAMRDPSSSEDPMEYRVLIGPPTNRIIF	352	
QY	1299	TVAVYFKGRIRIGCGKPSIQOAEHMA--AMDALEKYNPPOM-AHQKRIIGKRYRELKE	1354	
DB	353	KIAVYVYKGRILASPAESNVHKAELRYAELALANLIESMSPSKKAKANNMRRRLLEDQTS	412	

RESULT 2
 T21420
 hypothetical protein F26E4.10 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C/Accession: T21420
 R/Lighting: J.
 submitted to the EMBL Data Library, October 1996
 A/Reference number: Z19419
 A/Accession: T21420
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-682 <MW>
 A/Cross-references: UNIPARC:UPI0000179272; EMBL:Z81070; PIDD: CAB03006.1; GSPDB: GN00019,
 C/Experimental source: clone F26E4
 C/Genetics:
 A/Gene: CESP:F26E4.10
 A/Map position: 1
 A/Introns: 119/3; 226/2; 303/2; 377/3
 A/Superfamily: *Caenorhabditis elegans* hypothetical protein F26E4.10

	Query Match	Score	DB 2;	Length	682;
Best Local Similarly	7.4%;	558;			
Matches 178; Conservative	24.1%;	Pred. No. 4.4e-22;			
	16;	Mismatches 296;		Indels 140;	Gaps 23;

QY 277 RHRSYERSENERERHRHRDNRSSLSERSYKKEKRGSGSGYGLSVPRPAGCPPELP-C 335
Db 14 KHKARRRKYQKEYOERHKEEMOOLGRRONQSTSSAPDVTYEKIPEYTESALPFG 73

```

QY      336 EIINKTDS-----WAPPLEIVNHRSPSREKKRARWEEKDRWSDNQSGKDNKNTYSIK 388
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB      74 DSPRTEKQYETNYMIDPP--VSTSAELK-----SNRWIK 110

```

```
Qy      389 EKEPEETMPDKNEEEEEELK-----PWIRCTHS          418  
       :|::||:           ||  
Db     111 AEEAEKYMMTKAKSTTSKILOFQTILLETYVTKRRIQADVYIIHPCHSMGKRKTPrOK 170
```

```

Qy      419  ---ENYSSDPMQVDSDTVVGTSRLRLDYDKFEELGSRQKAKARPP--WEPPPTYKL 473
          ||: || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      171  GGDEFTASDVSDSDS-----QDEASTSEPTNRQAPADKT 208

```

```

09 474 DEDLESSSECESEDSTG-----SSDSEVFDVIAIEIKKKAHPDRLH 519
      |  ||  ||  :
Db 209 GE-----VKDEKQTCNRNQKAKRLNFEKEKQITLLKKGIDRKTTHPNIH 258

```

```

Oy      520 DELWYNDPGOMNDGPELCKCSAKARRTGRHSIYPGEBAIKPRPMTNNAAGRLTHY--RIT 577
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      259 PDISFNEKGLNGEGPECRCEPEIKTQGLKHGYVAGEDKAIDCK--KSGNGENLHYTLRVT 316

```

```

Oy      578  VSPPINFLTRPTVIEYDDHYIIFEGGSMFAAPL-----TNIPLCXIRFNIDYTIHIE 633
      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      317  PLPSENQL--YRTHMAINGEEFEGGSLITAPLPDCMTRAPICK---YSMDYEFQIVE 371

```

```

Qy      634  E M P P - E N F C V K G L E L F S L F L R D I L E Y D M N I K G L F E D S P P C C P R T H E M P R V R F L P D G 6922
      ||| | : : : : : : : : | | | | | :
Db      372  E F M P E C F D P E D C D M L F E Y I F H E I F E M L D F E L R P K H I P S D V E S C P M I H M P R F V Q - T K D D 4300

```

```

693 GKEVLSMHOILLYL-RCSKALVPBEERIANLQWEELEWOKKAAEECKGMIVNPGTIPSS 751
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 431 LVQLMSSKTVLAFYFTSKSSSEMSPEVNRCLDAQIDQFTRNTSKHKQSIIVNTKFKRPSA 490

```

```
Oy      752 VRIDOLDREQENPDVITFPILIVHFGIRPAQLSYAGDPOYOKLWMSYKLRHLANSRKYK 8111
```

```
Db      491 IRADFERDEKEKV-----YVNHNATRAQTYYAISLEPRTAFLEKTLMNGIQ-EKSSGSY 5455
```

```

0y      812 QTDKQKLAQREBEALQKIRQKQKTMREVTVLSSQGFNKTGISDVCHQAMMLPVLTHHIR 8711
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      546 NKDFEK---TKNELEHLKREKRSARNKTLREPVAQGFJETGLKPDVAAHVMMTILACHHIR 6022

```

Qy 872 YHOCIMHLDKIKGYTFQDRCLLQIATMTHPSHHILNFGKNPDPARNSLSNCGIOPKYGDRK 931

Db 603 YNFSLDVPEEIEIYKKFNDRVIELALMHSFKSHGYPIIDHYKMIINCGYRR-KYG--- 658

```
QY      932 VHHMHRKKGGINTLINIMSR 951
          :|:|      :  : :
Db      659 AEDKREKKRGSRLKJLKFDDK 678
```

RESULT 3
SS7447

N:Alternate names: HPBRII-4 protein
C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Oct-2004

R. Fleischhauer, K.L.
submitted to the EMBL Data Library, June 1992
A/Reference number: S57447

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-51 <FLB>

A:Accession: S57489
A:Status: preliminary
A:Molecule type: mRNA

A:Cross-references: UNIPARC:UP1000006D566; EMBL:X65337; NID:S871298; PIDN:CAA4752.1; PIR:G10881; C:Genetics: 231/3; A:Introns: 231/3

Query Match	4.2%	Score 316; DB 2;	Length 551;
Best Local Similarity	28.0%;	Pred. No. 2.1e-09;	

```

19 PRGRGCHGARPSAPSFRRPQNTLRLLHPQQRPVQYQYEBBPSAPTTFSNSDAPNF----LPP 74
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

75 RPDVFPPPPMPSAQGLPPCPRI RPP-----PPNHQNRHPPV-----DPC--FPP 11
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

120 MPPEMPCNNPVPGPAPGQ-----GTFP-----FMMPBSMPHPP--PPV 15

```

QY      160 MPQQVNYGYPPGYSHHNFPSPSFNSSQNPPSSLPSANNSSGP---HFRHLPRX----- 210
          ||| : ||||| : |:::~| ::|||
          +--+ -+-- --+-+--+ ---+----+-----+-----+-----+

```

```
QY      211 PLPKASERSPERLKGHYDDHRHDS-----                23
```

238 -----HG-RGERHRSLDRERGRSPRRRQDSRYSPDYDRGRTP 275

276 SRHRSYERSRERERHRHNDNRBSLSLERSYKEVK 312

RESULT 4

ribonuclease III (EC 3.1.26.3) - *Bacillus subtilis*
N/Alternate names: RNase D; RNase O
C:Species: *Bacillus subtilis*

R. Kunst, F. Gossawara, N. J. Mossner, I. Albertini, A. M. Alloni, G. Azavedo, V. Berterio, C. Bron, S. Broulliet, S. Brusch, C. V. Caldwell, B. Canano, V. Carceri, N. M. Chou, A. C. Enrich, S. D. Emmerson, P. T. Entian, K. D. Errington, J. J. Fabret, C. Ferrari, E. A. J. Enrich, S. D. Emmerson, P. T. Entian, K. D. Errington, J. J. Fabret, C. Ferrari, E.

Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fume, S.; Galizzi, A.; Gallen
Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koeber, P.; Koningsstein, G.; Krog, S.; Kumano, M.; Kurita, K.; Lapilus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maseel
Y.M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Rothe, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schlicht, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Serot
Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Wintere, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; PMID:98044033; PMID:9384377
A:Accession: B63693
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-249 <KUN>
A:Cross-references: UNIPROT: P51833; UNIPARC: UP1000060303; GB: Z99112; GB: AL009126; NID: G
A:Experimental source: strain 168
A:Oguro, A.; Kakeshita, H.; Takamatsu, H.; Nakamura, K.; Yamane, K.
Gene 172, 17-24, 1996
A:Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit, on
A:Reference number: J04819; PMID:96257247; PMID:8654983
A:Accession: J04821
A:Molecule type: DNA
A:Residues: 1-78, 'P', 80-137, 'K', 139-170, 'P', 172-242, 'E', 244-249, 'INPPYDSGGFQVCRLL' <OGU
A:Cross-references: UNIPARC: UP1000017592; DBJ: D64116; NID: G1389548
A:Note: the sequence in Genbank entry D64116, release 106.0, (PID:G1389549) has 138-Glu
C:Genetics:
A:Gene: rncs; srb
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
C:Keywords: hydrolase
F.172-245/Domain: double-stranded RNA-binding repeat homology <DSR>
Query Match 4.2%; Score 315; DB 2; Length 249;
Best Local Similarity 34.7%; Pred. No. 8.7e-10;
Matches 85; Conservative 46; Mismatches 96; Indels 18; Gaps 7;
Qy 1104 LOKLTFEFAIGVIFTHVRLAARAFLLRTVGFMHLTG--NORMEFLGDSIMQVATLEY 1161
Db 17 VQPFQFQRIHVHFNQEKLYQAFHSSVYNEHKKYEDNEHLEFLGDAVLELTSKF 76
Qy 1162 LFIHPDHGHGHTLRLSSLVNNRTOAKVAEELGMOEYAITNDKTRPVG-LRTKTLADL 1220
Db 77 LFAKPRMSEGLTLTKRAIVCEPSLVSLAHELSEFDVLLGKGEEMTGGRKRPALLADV 136
Qy 1221 LESFLAALTYDYLEVHTFMVAVCFPRRLKEFLINDQNDPVSQLOCCCLTRTEGK--- 1277
Db 137 FEAFGLATLDGGLFVSEFLLKVVYFPRKINDGAFSHVM-DIFSQLOE---YVORDKGSLL 192
Qy 1278 EEDIDPLVLTQVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEKYNFPQM 1337
Db 193 E-----YKTSNEKGPANHEFAIVSLKGPGLGVGNGSKAEHQAADALAKL---OK 244
Qy 1338 AHQKR 1342
Db 245 HTTKQ 249
RESULT 5
AE1300
ribonuclease III homolog rncs [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1300
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Kars, U.
Science 294, 849-853, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunat, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schueter, T.; Simeos, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; PMID:21537279; PMID:11676669

A:Accession: AE1300
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <GLA>
A:Cross-references: UNIPROT: O8Y691; UNIPARC: UP1000005520B; GB: NC_003210; PIDN: CAC99883.1;
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: rncs
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
Query Match 4.1%; Score 307.5; DB 2; Length 229;
Best Local Similarity 33.0%; Pred. No. 2e-09;
Matches 76; Conservative 51; Mismatches 86; Indels 17; Gaps 7;
Qy 1109 EFEAIGVIFTHVRLAARAFLLRTVGFMHL--TLGHNMREFLGSIMQVATLEYFIHF 1166
Db 6 ELQESVGFDFQVVELLQKQAFTHSSVYNEHRRNVNDNELEFLGDAVLELTVSDYLFNXY 65
Qy 1167 PPHHGHGHTLRLSSLVNNRTOAKVAEELGMOEYAITNDKTRPVGLRTK--TLADLSEF 1224
Db 66 PMAEGHMTKRAAIVCEPSLVEFAVAFSKY-VRLGGEKAGGRTRPALLADVSESF 124
Qy 1225 IAAITDLDLEVHTFMVAVCFPRRLK--EFLINDQNDPVSQLOCCCLTRTEGKPDIP 1282
Db 125 ICAIYLDNGIDKRVTFLEERVTFPRKIDAGAYLQTVDY--KTLDQETV-----QRDRDV 175
Qy 1283 L-YKTLQTVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEK 1331
Db 176 IEYDILGERTGPANHKAFAQVAVVNGVVLGKSGRTKQAEQSAADAPRANK 225
RESULT 6
AE1672
ribonuclease III homolog rncs [imported] - *Listeria innocua* (strain C1p11262)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1672
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauget, O.; Entian, K.D.; Fsihl, H.;
D.; Jones, L.M.; Kars, U.
Science 294, 849-853, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunat, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schueter, T.; Simeos, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; PMID:21537279; PMID:11676669
A:Accession: AE1672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <GLA>
A:Cross-references: UNIPROT: O92AK3; UNIPARC: UP10000134353; GB: AL592022; PIDN: CAC97149.1;
A:Experimental source: strain C1p11262
C:Genetics:
A:Gene: rncs
C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
Query Match 4.0%; Score 300; DB 2; Length 229;
Best Local Similarity 31.8%; Pred. No. 4.9e-09;
Matches 76; Conservative 53; Mismatches 88; Indels 22; Gaps 8;
Qy 1109 EFEAIGVIFTHVRLAARAFLLRTVGFMHL--TLGHNMREFLGSIMQVATLEYFIHF 1166
Db 6 ELQESVGFDFQVVELLQKQAFTHSSVYNEHRRNVNDNELEFLGDAVLELTVSDYLFNXY 65
Qy 1167 PPHHGHGHTLRLSSLVNNRTOAKVAEELGMOEYAITNDKTRPVGLRTK--TLADLSEF 1224
Db 66 PMAEGHMTKRAAIVCEPSLVEFAVAFSKY-VRLGGEKAGGRTRPALLADVSESF 124
Qy 1225 IAAITDLDLEVHTFMVAVCFPRRLK--EFLINDQNDPVSQLOCCCLTRTEGKPDIP 1282
Db 125 ICAIYLDNGIDKRVTFLEERVTFPRKIDAGAYLQTVDY--KTLDQETV-----QRDRDV 175
Qy 1283 L-YKTLQTVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEKYNFPQMAQ 1340
Db 1283 L-YKTLQTVGPBHARTYVAVYFKGERIGCGKSPSIQQAEMGAMDALEKYNFPQMAQ 1340

Db 176 IEYDLIGETGPAHKAFAQAQVIVNGOVLGSGSGRTKKAQEGSAQAFAIN-----QUTHR 229

RESULT 7

742727

proliferation potential-related protein - mouse
C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

C:Accession: T42727

R:Mitte, M.M.; Scott, R.E.
submitted to the EMBL Data Library, November 1998

A:Reference number: Z22246

A:Accession: T42727

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1560 <M17>

A:Cross-references: UNIPROT:P97868; UNIPARC:UPI0000028767; EMBL:U83913; NID:G9358884; PI

A:Experimental source: strain Balb/C

C:Genetics:

A:Gene: P2P-R

C:Function:
A:Description: Involved in hnRNP association and Rb1 binding

F:57-107/Domain: RING finger homology <RRN>

Query Match 3.9%; Score 293; DB 2; Length 1560;
Best Local Similarity 23.8%; Pred. No. 1.3e-07;
Matches 151; Conservative 72; Mismatches 200; Indels 212; Gaps 34;

QY 25 HBARSA--PSFRPQLRLHQQPPVOGYEPPSA-----PSTFTSNAPANFPLP 73

Db 296 NARPEGGPGWHSN-KLGYLVSPPOQIRGERSSTYRINGRHHSRSQTIQDS-IP 353

QY 74 PPDPFVPPPPPPSAQGLPCCPIRPPPNQMRHPPFVPPPCFPMPPMPCNNPPVP 133

Db 354 ARPVFPVPVPP-----PLYPPP-----HTLPLP-----PQVLPQGSPPQP--P 391

QY 134 GAPPGCGTFPPMPPSPMPP-----PPVMPQOVN----- 166

Db 392 GQPPTAG---YSPPPGPPAPANISTACSPGVPTAHSNMTPTTQAPLISBEFFREOK 448

QY 167 -----QVP---PGVHNHFPSPFNSPQNNPSPLPSANNSSPHFRLLPYPLPKAS 217

Db 449 RIKESEKFPYSGSSYSRSSTYDSKRSSTRSRSYSRSKSHSSTSPYP----- 502

QY 218 ERSRPERLKHYPDHRDHSHGGERHRSLDREGRSPDRRDSRYSDYD-RGRTPS 276

Db 503 -RRGRCKSNY---RRSRSHG---YHRS-----RRSPFYRRYHRSRSPQAFRGQST 550

QY 277 RHRYSYRSRERER-RRHR-----DNRSPLEERSYKKEYR 313

Db 551 K-RNVPRG-ETEREYFNRYREVPPPYDIKAYGRSVDFRDPFEKERYREMERKYREMYEK 608

QY 314 SGRSYLVSVPPRAGCTPPLPGIINKNTSMAP-----PLEIVNHRSPREKKRAREEEK 369

Db 609 YKGVAVGAQPRSA-----NRDPSPERLLPLNIN--SPFTRGRREDYAAQ 655

QY 370 DRWSDN-----QSSGDKNTYSIKEPEETMPD-----KNE---EE 403

Db 656 SHRNRLGNYPKLSTRSHNAKONPKSKESAVPDGKGNKHKHKKRKNRKEKGE 715

QY 404 EELLKPVI---RCHSENT-----YSSDMOVGDSTVVGTSR 440

Db 716 SESFLNPBELLGPKRCKRGSSGIDETKTDTLFFVPPSHEDATPVVDEBMD--ASITPKSVS 773

QY 441 LADLVKFEELGSRBEKAKA-----APPPEPKTULDELESSESE---CES 487

Db 774 DDKRERKDPKVKSDTKRKSDGSAATAKONVLKPSGQEKVDGUREKSPSEPLPKA 833

QY 488 DEDST-----CSSSDSEVPDVIATIKRKAH 514

Db 834 KEATKIDSVKPSSSSQKDEKVTGT-----PKAH 863

RESULT 8

B82073

ribonuclease III VC2461 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: B82073

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardoon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bess, S.; Qin, H.; Dragol, I.; Sellers, P.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: B82073

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-225 <HE1>

A:Cross-references: UNIPROT:Q9KPB2; UNIPARC:UPI0000134369; GB:AE004316; GB:AE003852; NID:

A:Experimental source: serogroup O1, strain N16961, biotype El Tor

C:Genetics:

A:Gene: VC2461

A:Map position: 1

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 285.5; DB 2; Length 225;
Best Local Similarity 32.2%; Pred. No. 2.8e-08;
Matches 75; Conservative 42; Mismatches 101; Indels 15; Gaps 5;

QY 1101 SPVLQKTEFEBAIGVIFTHVRLARAFTLRTVGFNNHLLTGHNQMEFLGDSIMOLVATE 1160

Db 2 TPNMKLT---SKGYTKETELNLNLTHRSANGK-----HNERLEFLGDSILSFVID 53

QY 1161 YLFHPDHGHGLTLRSSLVNNRTQAKVAEELGMOEATITNDTKRPGVGLRTKT-LAD 1219

Db 54 ELYRRFPKYNEDGMSRMATVIRGNTLAELGREFLDGVDLKLGPBELKSGGFRRSIDLD 113

QY 1220 LLESFIALYTDKOLEVYHTEFNVCFFPRLKEFLINQMDNDPKSLOQCCLLTRTEGKEP 1279

Db 114 AVEALITGALYLDSDLETARSTVLEWYHGLEIKFGASGDKTKRIQER-----LQGRKK 168

QY 1280 DIPLYTKQTVPSPHARTYTVAVVEFG-ERIGCGGPSIOQAEAGMADALEK 1331

Db 169 PLPVTVTNKGEANQERTVACVAGMDTPVIGKTSRKRAEQAAAEALRQ 221

RESULT 9

A83961

ribonuclease III rncs [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: A83961

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83960; MUID:20512582; PMID:11058132

A:Accession: A83961

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <STO>

A:Cross-references: UNIPROT:Q9KA05; UNIPARC:UPI0000134340; GB:AP001515; GB:BA000004; NID:

A:Experimental source: strain C-125

C:Genetics:

A:Gene: rncs

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 282; DB 2; Length 263;
Best Local Similarity 30.9%; Pred. No. 5.2e-08;
Matches 77; Conservative 50; Mismatches 110; Indels 12; Gaps 5;

QY 1084 HPLQLOEPNTDQLETSPVLQKTEFEBAIGVIFTHVRLARAFTLRTVGFNNH-LTIG 1141

Db 15 HSERRRP---KRLTLTKAQGMFDELLRTNL/LTFANKKLLVOAFTHSYVNEHRIQSGK 71

QY 1142 HNQRMEFLGDSIMOLVATEYLFIFHPDHGHGLTLRSSLVNNRTQAKVAEELGMOEYAI 1201

```
Db DNERLEFLDPAVLELAVSOYLVAPEOMSEBGMETKLRASIVCEPSLAQIAEELHPELVL 131
Oy 1202 TNDKTRPVG-LRTKTLADLLESFIALATYDKOLEVYHFMVNCFFPRLEKFLINDQMD 1260
Db 132 LGKGEEMTGRRKRPALLAVFESFVALYDQGMADVLFLERTTYPKISEGAFSH-MMD 190
Oy 1261 PKSOLOCCCLTLRTGSKERDIPLYKTLQTVGSPSHARTYVAVYFKGERIGCGKSPSIOQA 1320
Db 191 FFSQLOEF-----IQDNDIGHIHYEIVQERGRPHNNEFSEVVLNNETLIGVGRSKXA 245
Oy 1321 EKGAMDAL 1329
Db 246 EGHAAQOAL 254

RESULT 10
ribonuclease III (EC 3.1.26.3) [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: F98012
R/Host: J.A.; Albom Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Author: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: F98012
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-232 <KUR>
A/Cross-references: UNIPROT:Q8DPB8; UNIPARC:UPI00000835D9; GB:AE007317; PIDD:AAK9930.1;
C/Genetics:
A/Gene: rncS
A/Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 281.5; DB 2; Length 232;
Best Local Similarity 32.9%; Pred. No. 4.7e-08;
Matches 79; Conservative 47; Mismatches 91; Indels 23; Gaps 9;

Oy 1099 ETSPLVQLKTEFEAIGVIFTHVRLARAFTLRTVGFNH--LTLGHNMREFLGSIMOL 1156
Db 3 ELQTVAKNFEI-----FADKKLETAFTHTSVANERHLKISNELEFLGDAVLQI 56
Oy 1157 VATEYLFIHPDHHGHLTLRSSLVNNRTQAKVAELQMEYAITNDKTRPVGLRTK- 1215
Db 57 LISEVLYKKYPPKPEGDLSTLRAMIVREESLAGFARDQPDF-IKLGKEEKSGGRND 115
Oy 1216 -TLADLESFIALATYDKOLEVYHFMVNCFFPRLK--EFLINDQMDPKSOLOCCCLTL 1272
Db 116 TLIGDAFEAFGLALDKOVAKVEFIQVMIPKVAEGEMTIDY--KTHIQE--LL 169
Oy 1273 RTEGKEPDIPL-YKTLQTVGSPSHARTYVAVYFKGERIGCGKSPSIOQAEKGAMDALEK 1331
Db 170 QVNG--DVAIRQVISETGPAHDKVPDEVVLVEGSGIGGGRSKLAEGEAKAVAEK 226

RESULT 11
H95144
ribonuclease III [imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C/Accession: H95144
R/Host: J.A.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
neon, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Author: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: H95144
```

```
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-232 <KUR>
A/Cross-references: UNIPROT:Q970G6; UNIPARC:UPI0000051783; GB:AE005672; PIDD:AAK75353.1;
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP1248
A/Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.7%; Score 279.5; DB 2; Length 232;
Best Local Similarity 32.5%; Pred. No. 6e-08;
Matches 75; Conservative 48; Mismatches 91; Indels 17; Gaps 8;

Oy 1108 TEFEAIGVIFTHVRLARAFTLRTVGFNH--LTLGHNMREFLGSIMQVATEYLFIH 1165
Db 6 TVLKNHFAIEFADKKLETAFTHTSVANERHLKISNERLEFLGDAVQLISEFLYKK 65
Oy 1166 PPDHNEGHLTLRSSLVNNRTQAKVAELQMEYAITNDKTRPVGLRTK--TLADLES 1223
Db 66 YPKPEGDLSTLRAMIVREESLAGFARDQPDF-IKLGKEEKSGGRNDTLIGDAFEA 124
Oy 1224 FLAALYTDKOLEVYHFMVNCFFPRLK--EFLINDQMDPKSOLOCCCLTLRTGKEPDI 1281
Db 125 FGLALLDDKOVAKVEFIQVMIPKVAEGEFMTIDY--KTHIQE--LLQVNG--DV 175
Oy 1282 PL-YKTLQTVGSPSHARTYVAVYFKGERIGCGKSPSIOQAEKGAMDALEK 1331
Db 176 AIRQVISETGPAHDKVPDEVVLVEGSGIGGGRSKLAEGEAKAVAEK 226

RESULT 12
A57640
reclinoblastoma binding protein RBQ-1 - human
C/Species: Homo sapiens (man)
C/Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C/Accession: A57640
R/Host: Sakai, Y.; Saijo, M.; Coelho, K.; Kishino, T.; Nikiawa, N.; Taya, Y.
Genomics 30, 98-101, 1995
A/Title: cDNA sequence and chromosomal localization of a novel human protein, RBQ-1 (RBB1
A/Reference number: A57640; MUID:96129310; PMID:8595913
A/Accession: A57640
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-948 <SAK>
A/Cross-references: UNIPROT:Q15290; UNIPARC:UPI000006DBD4; GB:X85133; NID:g728590; PIDD:
C/Genetics:
A/Gene: GDB:RBBP6
A/Cross-references: GDB:626076
A/Map position: 16p12-16p11.2
C/Keywords: tandem repeat; zinc
F;79-129/Domain: RING finger homology <RNG>

Query Match 3.7%; Score 276; DB 2; Length 948;
Best Local Similarity 23.0%; Pred. No. 5.5e-07;
Matches 149; Conservative 61; Mismatches 208; Indels 230; Gaps 31;

Oy 27 ARPSA--PSFRQNLRLHPQRPVQYQYBPSPA-----PSRTFSNPAFNPFLPR 75
Db 320 ARPGGRPGMEHSN-KLGLVSPPOQIRGERSCYRSIRGRHHSRSQRTQGS-LPAT 377
Oy 76 PPFVPPPPMPSAQGLPLPCPIRPPFPMHQMHPVPVPCFPPMPDPMPCPNNPVPFGA 135
Db 378 PVPVPVPPPP-----PLVPPPP-----HTLPLP--PGVPPPGQSPGF--PQ 415
Oy 136 PPGQGTFFPMMPSPMHPHPPPPPVFPQVN-----YQYPP----- 170
Db 416 PPAPG--YSVPPEPP-PAPNLSITPWSVSGVQTASHNTIPTQAPPLPSREFFYEQR 471
Oy 171 -----GSHHNPSPSFS 184
Db 472 LKEEEKKSLDEFNTPDKALMEYKKIKERRRSRSKSPSGSYRSSTTYSKSS 531
Oy 185 PQNNPSSFLPSANNSSPHFRHLPPYPLPKAPSERRSPRLKHYDDHRRHDSHGGERH 244
```

Dh	532	GSFTRSRYSRFSFSHSHSRYSRSPYP-----RRGRGKS RNY-----RSRSRSHG-----YH	578
QY	245	RSILDRRRGRGSPDRRQDSRYRSDYD-RCGRPSRHRHRYERSREER-----	289
Dh	579	RS-----RSRPPRYRHSRSRSPOAFRGSPKNRNPQGETEYENRNYREVPYPDMK	633
QY	290	-----RRHHDRNRSPSLERSYKKEKRSGRSGYGLSVPEPAGTPELPGC	336
Dh	634	AYGRASVDPRDPFEKERYE-----WERKRYEYKCYKGYAAGAPRBSA-----	679
QY	337	LIKNTDWP-----PLEIVNHSRSPREKKARWEEKDRWSDNGSG-----KD	381
Dh	680	---NRKNSPFRFLPLNIRN--SPFTRGRRDYVGGGSHSRNIGSNYPKLSARDGHNQ	734
QY	382	KNYTSIKEKEBEETMPD-----KNEEBEELKPVWIRCHSENYSSDPW--	427
Dh	735	KDNTKSKKESENA PGDQGNKHKKHKKRRKRKGESEGLNPELLETSRK-----SREPTGV	790
QY	428	-DQVGDSVTVGTSR-----LADLYDKFEEF--LGRQEKAKKARPMEPPTLDDLE	478
Dh	791	EENKTDLSFLVPSRQDADTPVRD--EPMDAESITTEKVSSEKDKRRR---DKFKAGDKTKR	845
QY	479	SSSKSEGESDEDSCTSSSSDSE-----VFDVIAEIKKKKAPDRLL	518
Dh	846	KNDGSAVSCKENIVKPAKPGQEKVDGVRDLLDNLQLKPKRRLRLRL	893

THIS PAGE BLANK (USPTO)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:32:44 ; Search time 287 Seconds
(Without alignments)
3377.686 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMQNTCHRMSPHGRGCRP.....MRMRSHQSRPDEDEDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniProt_sprot.*

2: uniProt_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7471	99.6	1374	1 RNC_HUMAN	Q9NTR4 homo sapien
2	7160	95.5	1373	2 Q5H2J0_MOUSE	Q5H2J0 mus musculu
3	6380	85.1	1231	2 Q6PAJ6_MOUSE	Q6PAJ6 mus musculu
4	6344.5	84.6	1236	2 Q5ZIR3_CHICK	Q5ZIR3 gallus gall
5	5217	69.6	1021	2 Q80Z69_MOUSE	Q80Z69 mus musculu
6	4982	66.4	942	2 Q6PF88_MOUSE	Q6PF88 mus musculu
7	4121	54.9	780	2 Q7TW18_MOUSE	Q7TW18 mus musculu
8	3767	50.2	1041	2 Q4SKM1_TETNG	Q4SKM1 tetraodon n
9	3696	49.3	705	2 Q80UJ5_MOUSE	Q80UJ5 mus musculu
10	2963	39.5	1327	2 Q9XN5_DROME	Q9XN5 drosophila
11	2881.5	38.4	979	2 Q5TOK8_ANOGA	Q5TOK8 anopheles g
12	2825	37.7	941	2 Q8CJ74_MOUSE	Q8CJ74 mus musculu
13	2811.5	37.5	1071	2 Q960Y4_DROME	Q960Y4 drosophila
14	2683.5	35.8	836	2 Q7PNE9_ANOGA	Q7PNE9 anopheles g
15	2298	30.6	432	2 Q5RFF9_HUMAN	Q5RFF9 homo sapien
16	1480	19.7	288	2 Q67B98_BAT	Q67B98 ratius norv
17	1473.5	19.6	1061	2 Q61XX9_CABER	Q61XX9 caenorhabdi
18	1448	19.3	1086	1 RNC_CABER	Q01326 caenorhabdi
19	1154.5	15.4	290	2 Q4SKM3_TETNG	Q4SKM3 tetraodon n
20	1138.5	15.2	412	2 Q569M4_XENLA	Q569M4 xenopus lae
21	711	9.5	134	2 Q9CTG2_MOUSE	Q9CTG2 mus musculu
22	344.5	4.6	220	2 Q82ZG1_ENTRA	Q82ZG1 enterococcu
23	340.5	4.5	1150	2 Q81MK6_DROME	Q81MK6 drosophila
24	340	4.5	846	2 Q8MRP6_DROME	Q8MRP6 drosophila
25	340	4.5	1109	2 Q9VAV4_DROME	Q9VAV4 drosophila
26	339	4.5	1389	2 Q5RFS1_DICDI	Q5RFS1 dictyosteli
27	331	4.4	1838	2 Q6BCJ9_TETTR	Q6BCJ9 tetrahymena
28	329.5	4.4	2869	2 Q6OPD2_CABER	Q6OPD2 caenorhabdi
29	328	4.4	1022	2 Q4SRX9_TETNG	Q4SRX9 tetraodon n
30	325	4.3	551	2 Q8BX18_MOUSE	Q8BX18 mus musculu
31	321	4.3	249	2 Q65J05_BACUD	Q65J05 bacillus

32	321	4.3	551	2 Q8BX86_MOUSE	Q8BX86 mus musculu
33	321	4.3	551	2 Q6NVF9_MOUSE	Q6NVF9 mus musculu
34	319	4.3	825	2 Q7S9H3_NEUCR	Q7S9H3 neuropept
35	317.5	4.2	552	2 Q5NVH8_PONPY	Q5NVH8 ponopygma
36	316	4.2	551	2 Q5IRK1_HUMAN	Q5IRK1 homo sapien
37	316	4.2	551	2 Q16630_HUMAN	Q16630 homo sapien
38	315	4.2	249	1 RNC_BACSU	Q51833 bacillus su
39	311.5	4.2	528	2 Q4SQ65_TETNG	Q4SQ65 tetraodon n
40	310	4.1	548	2 Q6DDW4_XENLA	Q6DDW4 xenopus lae
41	310	4.1	551	2 Q5ZL34_CHICK	Q5ZL34 gallus gall
42	309.5	4.1	1604	2 Q7X014_ORYSA	Q7X014 oryza sativ
43	308.5	4.1	245	1 RNC_BACCR	Q81978 bacillus ce
44	307.5	4.1	229	1 RNC_LISMO	Q81951 bacillus mo
45	307.5	4.1	245	1 RNC_BACAN	Q81W18 bacillus an

ALIGNMENTS

RESULT 1
ID RNC_HUMAN STANDARD: PRT: 1374 AA.
AC Q9NTR4; Q7Z5V2; Q86YH0; Q9NM73; Q9YZV9; Q9Y4T0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III) (Drosophila) (p241).
GN Name=RNASE3L; Synonym=RN3;
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=20538440; PubMed=10948199; DOI=10.1074/jbc.M005494200;
RA Wu H., Xu H., Miraglia L.J., Crooke S.T.;
RT "Human RNase III is a 160-kDa protein involved in preribosomal RNA processing".
RL J. Biol. Chem. 275:36957-36965(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 166-613 (ISOFORM 2).
RC TISSUE=Colon;
RX MEDLINE=20431278; PubMed=10976766; DOI=10.1023/A:1007177623283;
RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Maeda H.,
RA Sekine M., Obaishi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagatani K., Murakami K., Yasuda T., Iwayanagi T., Magatsuna M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamauchi M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Iihara S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hottu T.,
RA Kusano J., Kanohori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Mutsaers K., Yuuki H., Oshima A., Saeki N., Aoshima S.,
RA Yoshikawa Y., Matsumura H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukunumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami H., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,

RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togeishi T., Oyama M., Hata H., Matanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuhito Y., Yamashita K.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.,
 RA "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RT Nat. Genet. 36:40-45(2004).
 RN (4)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 3), AND NUCLEOTIDE
 RP SEQUENCE [LARGE SCALE MRNA] OF 653-1374 (ISOFORM 1).
 RC TISSUE=Cervix, and Skin;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE OF 706-1374.
 RC TISSUE=Aorta;
 RA Wei J.Y., Ding J.F., Xiong H., Zhou Y., Iqbal C.C.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP FUNCTION.
 RX PubMed=14508493; DOI=10.1038/nature01957;
 RA Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P.,
 RA Radmark O., Kim S., Kim V.N.;
 RT "The nuclear RNase III Drosha initiates microRNA processing.";
 RL Nature 425:415-419(2003).
 CC -1- FUNCTION: Executes the initial step of microRNA (miRNA) processing
 CC in the nucleus, that is cleavage of pri-miRNA to release pre-
 CC miRNA. Involved in pre-miRNA processing. Cleaves double-strand RNA
 CC and does not cleave single-strand RNA.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC phosphonooxocetate.
 CC -1- SUBUNIT: Interacts with Spl.
 CC -1- INTERACTION:
 CC Q8WYQ5.DGCR8; NBExp=2; Inact=EBI-528367, EBI-528411;
 CC -1- SUBCELLULAR LOCATION: Nuclear. A fraction is translocated to the
 CC nucleolus during the S phase of the cell cycle.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q9NR4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9NR4-2; Sequence=VSP_005777;
 CC Name=3;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q9NR4-3; Sequence=VSP_012450, VSP_012451, VSP_012452,
 CC VSP_012453;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
 CC -1- SIMILARITY: Contains 2 RNase III domains.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a

CC frameshift in position 775.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, AF189011; AAF80558.1; -, mRNA.
 CC EMBL, AJ242976; CAB4513.1; -, mRNA.
 CC EMBL, AK001121; BAA9151.1; ALT INIT; mRNA.
 CC EMBL, BC041162; AAH41162.1; -, mRNA.
 CC EMBL, BC041003; AAH41003.1; -, mRNA.
 CC EMBL, AF116910; AAD29637.1; ALT_FRAME; mRNA.
 CC HSSP; 067082; 1JF2.
 CC Inact; Q9NR4-1;
 CC Ensembl; ENSG00000113360; Homo sapiens.
 CC MIM; 608828;
 CC InterPro; IPR001159; Dg_RNA_bd.
 CC InterPro; IPR000999; RNase_III.
 CC Pfam; PF00035; dsm; 1.
 CC Pfam; PF00636; Ribonuclease_3; 2.
 CC SMART; SM00358; DSRM; 1.
 CC SMART; SM00355; RIBOC; 2.
 CC PROSITE; PS00137; DS_RBD; 1.
 CC PROSITE; PS00137; RNASE_3_1; 2.
 CC PROSITE; PS00142; RNASE_3_2; 2.
 CC KW Nuclease; Repeat; Ribosome biogenesis; Hydrolyase; Nuclear protein;
 FT DOMAIN 876 1056
 FT DOMAIN 1107 1233
 FT DOMAIN 1260 1334
 FT COMPBIAS 1 212
 FT COMPBIAS 219 316
 FT VARSPLIC 285 353
 FT FT
 FT FT
 FT VARSPLIC 316 352
 FT FT
 FT VARSPLIC 353 353
 FT FT
 FT VARSPLIC 1198 1229
 FT FT
 FT FT
 FT VARSPLIC 1230 1374
 FT FT
 FT CONFLICT 166 174
 FT CONFLICT 612 612
 FT CONFLICT 1020 1020
 FT CONFLICT 1230 1230
 SQ SEQUENCE 1374 AA; 159316 MW; ED6FDEA09F3B8092 CRC64;
 Query Match 99.6%; Score 7471; DB 1; Length 1374;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1370; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MMQGNCHMSFHPGRCGRGGRGARGPAPSPFRQNLRLHPOQPPVQYQIEPPSPAPS 60
 DB 1 MMQGNCHMSFHPGRCGRGGRGARGPAPSPFRQNLRLHPOQPPVQYQIEPPSPAPS 60
 QY 61 TTFSNSPAPNPLPRPDPVFPFPPMPSPAGPLPCPIRPPFNHQRHPPVPCPPM 120
 DB 61 TTFSNSPAPNPLPRPDPVFPFPPMPSPAGPLPCPIRPPFNHQRHPPVPCPPM 120
 QY 121 PPMPCPNPVPVPGAPGCGTFFPMKPPSPMHPPPPPVMPQOVNYQYPPGYSHNPPPP 180
 DB 121 PPMPCPNPVPVPGAPGCGTFFPMKPPSPMHPPPPPVMPQOVNYQYPPGYSHNPPPP 180
 QY 181 SFNSFQNNSSFLPSANNSSPHFRHLPPYPLPKASERSRRLKHGYDDHRRDHSRGR 240
 DB 181 SFNSFQNNSSFLPSANNSSPHFRHLPPYPLPKASERSRRLKHGYDDHRRDHSRGR 240

Qy	241	GERHRSILDRERGRSGPDRRRODSRYRSYDNGRTSPRHSYERSGERERERHRRDRNRS	300
Db	241	GERHRSILDRERGRSGPDRRRODSRYRSYDNGRTSPRHSYERSGERERERHRRDRNRS	300
Qy	301	PSLERSYKKEKYSGRSGYGLSVPEPACTPELPEEIIKNDSMAPLLEIVNHSPSREK	360
Db	301	PSLERSYKKEKYSGRSGYGLSVPEPACTPELPEEIIKNDSMAPLLEIVNHSPSREK	360
Qy	361	KRAWEEKDKWSDWOSGDKXNYTSIEKEPEEETMPDKNEEBEELIKPWIRCTHSEN	420
Db	361	KRAWEEKDKWSDWOSGDKXNYTSIEKEPEEETMPDKNEEBEELIKPWIRCTHSEN	420
Qy	421	YSSDDPMQVDSITYVGTSTRLDLYDKFEEBEISGQEKAKAPRPEWPKTKLDELESS	480
Db	421	YSSDDPMQVDSITYVGTSTRLDLYDKFEEBEISGQEKAKAPRPEWPKTKLDELESS	480
Qy	481	SESSCEDSDSDTSSSDSEVFDVIAEIKRKAHPRDLHDELWYNDPCQMDGPLCKCSA	540
Db	481	SESSCEDSDSDTSSSDSEVFDVIAEIKRKAHPRDLHDELWYNDPCQMDGPLCKCSA	540
Qy	541	KARBTGISHSYPGEBEAIKPCRPMNNAGRLFHYAITYSPPTNFJLDRPTYAEYDHEXI	600
Db	541	KARBTGISHSYPGEBEAIKPCRPMNNAGRLFHYAITYSPPTNFJLDRPTYAEYDHEXI	600
Qy	601	PEGFSMFAHBLVTJNPLCKVIRFNIDYTHIEBEMPENFCVKELEFLPLFRDILELY	660
Db	601	PEGFSMFAHBLVTJNPLCKVIRFNIDYTHIEBEMPENFCVKELEFLPLFRDILELY	660
Qy	661	DWNLKGPLFEEDSPCCPRFHMPFRVRLPROGKEVLSMHQILLYLRLCSALYBEEBIA	720
Db	661	DWNLKGPLFEEDSPCCPRFHMPFRVRLPROGKEVLSMHQILLYLRLCSALYBEEBIA	720
Qy	721	NMLQMEELBEMQKVAEBECKGMITVNBGTBSSRIRIDOLRBNPVRITFPIIYHNGIRPA	780
Db	721	NMLQMEELBEMQKVAEBECKGMITVNBGTBSSRIRIDOLRBNPVRITFPIIYHNGIRPA	780
Qy	781	QLSYAGDPQOYOKLWKSYYKLRHLHLANSFKVQOTDKOKLAOREEALOKIRQNTMRRETV	840
Db	781	QLSYAGDPQOYOKLWKSYYKLRHLHLANSFKVQOTDKOKLAOREEALOKIRQNTMRRETV	840
Qy	841	ELSSQGFMTKGIIRSDVCQHAMLPLYLTHIRYHQCIMLIDYLGTFODRCLLOLMTHP	900
Db	841	ELSSQGFMTKGIIRSDVCQHAMLPLYLTHIRYHQCIMLIDYLGTFODRCLLOLMTHP	900
Qy	901	SHHLNPGNPDHANSLNSNGIRQPKYDGRVYHNMHMKKINTLINMISLGGDDPTPS	960
Db	901	SHHLNPGNPDHANSLNSNGIRQPKYDGRVYHNMHMKKINTLINMISLGGDDPTPS	960
Qy	961	RINNERLEFJGDNAVEFLTSVHLYYLPELSBEGGLATYTRAIYONQHILMIAKLLEDP	1020
Db	961	RINNERLEFJGDNAVEFLTSVHLYYLPELSBEGGLATYTRAIYONQHILMIAKLLEDR	1020
Qy	1021	FMLYAHGPDLCRESDLRHHAMANCEALIGAUYLEBSLEBAQOLFGRLLFNDPDLREVLIN	1080
Db	1021	FMLYAHGPDLCRESDLRHHAMANCEALIGAUYLEBSLEBAQOLFGRLLFNDPDLREVLIN	1080
Qy	1081	YPLHPLQLOEBNTRQOLIETSPVLOKLTREBEAIGVITTHVRLARAFATLTATVGNHILT	1140
Db	1081	YPLHPLQLOEBNTRQOLIETSPVLOKLTREBEAIGVITTHVRLARAFATLTATVGNHILT	1140
Qy	1141	GHNORMEFLGDSIMQVATEYLFHFPHHGHGHTLLRSSLVNNRTOAKVAEELGMOEYA	1200
Db	1141	GHNORMEFLGDSIMQVATEYLFHFPHHGHGHTLLRSSLVNNRTOAKVAEELGMOEYA	1200
Qy	1201	ITNDTKRPAVLARTKTLADLESFIALYTDKOLEVYTFPMNVGCFPPRLKFEIINOMDND	1260
Db	1201	ITNDTKRPAVLARTKTLADLESFIALYTDKOLEVYTFPMNVGCFPPRLKFEIINOMDND	1260
Qy	1261	PKSGLOOCCCLTLRIBGKEBDIPLKYTLQTVGSPHARTYTVAVYRGERIIGCGKSPSIOQA	1320
Db	1261	PKSGLOOCCCLTLRIBGKEBDIPLKYTLQTVGSPHARTYTVAVYRGERIIGCGKSPSIOQA	1320

QY	1371	MMSGAMDALEKXNPPOMAHQRKFRIGRKRYROGLKEMRWERREHOEPEDETEIDKK	1374
DB	1371	EMGAAMDALERTXNPPOMAHQKRIFERTKYRQELKENWREREHQBEPDETEDIKK	1374
RESULT 2			
OSHZJ0_MOUSE			
ID	OSHZJ0_MOUSE PRELIMINARY;	PRT; 1373 AA.	
AC	OSHZJ0;		
DT	10-MAY-2005 (TREMBLrel. 30, Created)		
DR	10-MAY-2005 (TREMBLrel. 30, Last sequence update)		
DE	10-MAY-2005 (TREMBLrel. 30, Last annotation update)		
GN	Nasens protein.		
OS	Name=Scn12; Synonym=Rnaesen;		
OC	Mus musculus (Mouse).		
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridea; Muridae; Murinae; Mus.		
CX	NCBI_TaxID=10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	STRAIN=C57BL/6; TISSUE=Brain;		
	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.W., Schuler G.D.,		
RA	Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hasteh P.,		
RA	Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udell T.B., Toshitsuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,		
RA	Rohak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Bikeshley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Gilwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skelton U., Smallus D.E.,		
RT	Scherer A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RL	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=C57BL/6; TISSUE=Brain;		
RG	NIH MC Project:		
RL	Submitted (UN-2005) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL, BC088999; AAC8999.1; - mRNA.		
DR	Ensembl; ENSMUSG00000212191; Mus musculus.		
DR	MGI; MGI:1261425; Etch12.		
DR	MGI; MGI:1261425; Nasens.		
DR	GO; GO:0005622; C:intracellular; IEA.		
DR	GO; GO:0003725; F:double-stranded RNA binding; IEA.		
DR	GO; GO:0004525; F:ribonuclease III activity; IEA.		
DR	GO; GO:0006396; P:RNA processing; IEA.		
DR	InterPro; IPR001159; De RNA bd.		
DR	InterPro; IPR000999; RNase_III.		
DR	pfam; PF00035; dnmr. 1.		
DR	pfam; PF00636; Ribonuclease_3; 2.		
DR	SMART; SMART0358; DSRM. 1.		
DR	SMART; SMART0358; RIBOC. 2.		
DR	PROSITE; PS01377; DS_RBD. 1.		
DR	PROSITE; PS00517; RNASE_3_1; 2.		
DR	PROSITE; PS01424; RNASE_3_2; 2.		
SO	SEQUENCE 1373 AA; EDIDBEE2FFD89A6B CRC64;		
Query Match	95.5%; Score 7160; DB 2; Length 1373;		
Best Local Similarity	95.3%; Pred. No. 0;		
Matches 1310, Conservative 32, Mismatches 30; Indels 2; Gaps 2;			
2 MGGNTCHMSFHPGRGCCPRGGHGAGPSAPSPFPONILRIIHPQPVPVOYEPPSAST 61			

Db 1 MCGNTCHRNSTYHGRGCGRGGHGARGBPAPAFRPNLRLHLQDPAPAOYQTEPPSPASBS 60
Qy 62 TSNBPAPNLPRLPDPFVFPFPPMPSAOGPLPCEIRPPFPNHQMRHPVPVPCPPMP 121
Db 61 SSSNQAPBFMPRPBFVFPFPPAPASAGPLPCCVPRPPYNNHQMRHFPVPCPPMP 120
Qy 122 PMPCCNPNDVGAAPGCGTFFPMMP PSMMPHPPPPVMPQOYNYQYPPGYSHANFPSPS 181
Db 121 PMPCCNPNDVGAAPGCGTFFPMMP PSMMPHPPPPVMPQOYNYQYPPGYSHANFPSPS 179
Qy 182 PMSFONNPSFLPSANSSSPHFRHL.PYPLPKABERSPERLKYDHRHSHGSG 241
Db 180 PMSYQNNSSFPSPYANSSSTPHFRHL.PYSLPQAKNERSPERLKYDHRHSHGSG 239
Qy 242 EHRHSLDRERGRSPDRRRQDSYRSDYRGRT -PSRHSYERSRERERHRHRDRRS 300
Db 240 EHRHSLDRERGRSPDRRRQDSYRSDYRGRT -PSRHSYERSRERERHRHRDRRS 299
Qy 301 PSLERSYKKEYYKSGSGSYGLSVPEPAGCTPELPGELIKNTDSMAPPLEIVNHRSPRK 360
Db 300 PSLERSYKKEYYKSGSGSYGLSVPEPAGCTPELPGELIKNTDSMAPPPENVNHRSRSPRK 359
Qy 361 KARPEEEDRWSDNOSGKOKKYTSIKKEPEPTMPDNKEESEEELPKPMVIRCHSHSN 420
Db 360 KARPEEEDRWSDNOSGKOKKYTSIKKEPEEVPPEKTEEESEEELPKPMVIRCHSHSN 419
Qy 421 YSSDPMDOGVSTVGTGRLRLDYKFEELGSRQEKAKAAPMPPEPKTKLDEDESS 480
Db 420 YSSDPMDOGVSTVGTGRLRLDYKFEELGSRQEKAKAAPMPPEPKTKLDEDESS 479
Qy 481 SESECESDSDSTCSSSDSEVPDVAIEIKRKKAHPRLHDELMYNDPGQMDGSLCKCSA 540
Db 480 SESECESDSDSTCSSSDSEVPDVAIEIKRKKAHPRLHDELMYNDPGQMDGSLCKCSA 539
Qy 541 KARRCTIRASTYRGEBAIKPCRPMTNNAGRLPHYR.LTVSPNMF.LDRCTVIEYDDEHYT 600
Db 540 KARRCTIRASTYRGEBAIKPCRPMTNNAGRLPHYR.LTVSPNMF.LDRCTVIEYDDEHYT 599
Qy 601 PEGFSMFAAPLTLNIPCKVIRENIDYTHFIEEMMPENFCVGLFELFLPRDILLEY 660
Db 600 PEGFSMFAAPLTLNIPCKVIRENIDYTHFIEEMMPENFCVGLFELFLPRDILLEY 659
Qy 661 DNMLKGPLFEDSPCCPRFHPMPRVRLPDGSEVLSNHQILLYLRCSKALVPEEETIA 720
Db 660 DNMLKGPLFEDSPCCPRFHPMPRVRLPDGSEVLSNHQILLYLRCSKALVPEEETIA 719
Qy 721 NMLQWELMOKYAEBCCKMIVNPGTKSSVRI.DQLDRBQEPVEYITPILVHFGIRPA 780
Db 720 NMLQWELMOKYAEBCCKMIVNPGTKSSVRI.DQLDRBQEPVEYITPILVHFGIRPA 779
Qy 781 QLSVAGDPQYQKLMSYVYKLRHLNLSAPKYQTDOKLAQREBALQKIRQKMTREEVY 840
Db 780 QLSVAGDPQYQKLMSYVYKLRHLNLSAPKYQTDOKLAQREBALQKIRQKMTREEVY 839
Qy 841 ELSSQGFMTGTGRSDVCOHAMMLPVLTTHIRYHQCMLMDKLGITYFORCLQLAMTHP 900
Db 840 ELSSQGFMTGTGRSDVCOHAMMLPVLTTHIRYHQCMLMDKLGITYFORCLQLAMTHP 899
Qy 901 SHHLNGMNPDRARNSLNCGRQPKYGRKXTHMMMRKGINTLNINSRLGODDPTGS 960
Db 900 SHHLNGMNPDRARNSLNCGRQPKYGRKXTHMMMRKGINTLNINSRLGODDPTGS 959
Qy 961 RINHNRLRFLGDAVVEFLTSVHL.YVLPSELSEGLATYRTAIVQNHMLAKKLELP 1020
Db 960 RINHNRLRFLGDAVVEFLTSVHL.YVLPSELSEGLATYRTAIVQNHMLAKKLELP 1019
Qy 1021 PMLYAHGPDLGRESDLRHMANCFEALIGA.VYLEGSLLEAKQJFGLFNDDPLREVMIN 1080
Db 1020 PMLYAHGPDLGRESDLRHMANCFEALIGA.VYLEGSLLEAKQJFGLFNDDPLREVMIN 1079
Qy 1081 YPLHPIQLOBPNTDROLIETSPVLOKLTREBEAIGVIFTHVRLLAAPFLIRTYGFNHLTL 1140
Db 1080 YPLHPIQLOBPNTDROLIETSPVLOKLTREBEAIGVIFTHVRLLAAPFLIRTYGFNHLTL 1139

Qy 1141 GHNQMEFLGDSIMOLVATEYTFIHPDHHGHLTLNLSLVNNRTQAKVAEELGQVEYA 1200
Db 1140 GHNQMEFLGDSIMOLVATEYTFIHPDHHGHLTLNLSLVNNRTQAKVAEELGQVEYA 1199
Qy 1201 ITNDKTKRPVGLRTKTLADLESFIALYTDQLEFYVHTFMVNCFFPRKEFILINDQND 1260
Db 1200 ITNDKTKRPVGLRTKTLADLESFIALYTDQLEFYVHTFMVNCFFPRKEFILINDQND 1259
Qy 1261 PXSQLOQCLTLRTEGKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCKGSPISQQA 1320
Db 1260 PXSQLOQCLTLRTEGKEBDIPLYKTLQTVGSHARTYVAVYFKGERIGCKGSPISQQA 1319
Qy 1321 EKGAMDALEKYNFQMAHQKRFGRKTRQELKEMRWEPENHOREPDEDEDIKK 1374
Db 1320 EKGAMDALEKYNFQMAHQKRFGRKTRQELKEMRWEPENHOREPDEDEDIKK 1373
RESULT 3
Q6PAU6_MOUSE
ID Q6PAU6_MOUSE PRELIMINARY; PRT; 1231 AA.
AC Q6PAU6;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Name=Etch12; Synonym=RNaseen;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain; DOI=10.1073/pnas.242603899;
RX MEDLINE=23388257; PubMed=12477932; PubMed=12477932; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.D., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (OCT-2003) to the EMBL/Genbank/DBD databases.
DR EMBL; BC060265; AAH60265.1; -; mRNA.
DR Ensembl; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etch12.
DR MGI; MGI:1261425; RNaseen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; De_RNA_bd.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dcrn; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.

DR SMART; SM00358; DSRM; 1.
 DR SMART; SM00535; RIBOC; 2.
 DR PROSITE; PS0137; DS_RBD; 1.
 DR PROSITE; PS00517; RNASE_3_1; 2.
 DR PROSITE; PS0142; RNASE_3_2; 2.
 FT NON_TER 1 1
 SQ SEQUENCE 1231 AA; 143573 MW; 3679FA9D87800151 CRC64;

Query Match 85.1%; Score 6380; DB 2; Length 1231;
 Best Local Similarity 96.1%; Pred. No. 2.9e-303;
 Matches 1184; Conservative 22; Mismatches 24; Indels 2; Gaps 2;

QY 144 FMWPEPMHPPEPPVWQVNYQYPPGYSHNFPPEPSFNSFQNNPSSFLPSANSSSPH 203
 1 FMVPEPMHPPEPPVWQVNYQYPPGYSHNFPPEPSFNSFQNNPSSFLPSANSSSPH 59

DB 204 FRLPYPLPKASERKSPRLKHYDDHRRDHSNGRGRHRLDRREGGRSDRRKODS 263
 60 FRLPYPLPKASERKSPRLKHYDDHRRDHSNGRGRHRLDRREGGRSDRRKODS 119

QY 264 RYASDVRDGRGRT-PSRRSYERSRERERERHRRDNRSPSLERSYKKEYSRSGRSTGLSY 322
 120 RYASDVRDGRGRTPPRRHRSYERSRERERERHRRDNRSPSLERSYKKEYSRSGRSTGLPY 179

QY 323 VEPACCTPELPGELIKNTDSMAPLEIVNHRSPSEKRRARWBEKDRWSDNQSSGKDK 382
 180 AEPACCTPELPGELIKNTDSMAPLEIVNHRSPSEKRRARWBEKDRWSDNQSSGKDK 239

QY 383 NNTSIEKEPEEEMPDONBEELKLPWIRCTHSENYSSDPNQVGDSTVVGTSRLR 442
 240 NNTSIEKEPEEEMPDONBEELKLPWIRCTHSENYSSDPNQVGDSTVVGTSRLR 299

QY 443 DLYDKEEELSGRQEKAKAAPWEPKTKLDEDLSSSECESEDSTCSSSDSEVF 502
 300 DLYDKEEELSGRQEKAKAAPWEPKTKLDEDLSSSECESEDSTCSSSDSEVF 359

QY 503 DVIAEIKRKKAKHDBRLHDELWYNDPGOMDGPLCKCSAKARRTGIRHSIYGEBAIKPCR 562
 360 DVIAEIKRKKAKHDBRLHDELWYNDPGOMDGPLCKCSAKARRTGIRHSIYGEBAIKPCR 419

QY 563 PMTNNAAGRLPHRYITVSPPTNFLTDRPTVIEYDHEIYFEGSGMFAHAPLTNIPCLKVIR 622
 420 PMTNNAAGRLPHRYITVSPPTNFLTDRPTVIEYDHEIYFEGSGMFAHAPLTNIPCLKVIR 479

QY 623 FNIDYIHFIEEMMPENFCVKGLEFSLFPRDILEYOMNKGPLFEDSPCCPFPHM 682
 480 FNIDYIHFIEEMMPENFCVKGLEFSLFPRDILEYOMNKGPLFEDSPCCPFPHM 539

QY 683 PRFVRFLPGGKKEVLSMHOILLYLRCSEKALVEEBEIANMLQWBELEWOKYAECKGMIV 742
 540 PRFVRFLPGGKKEVLSMHOILLYLRCSEKALVEEBEIANMLQWBELEWOKYAECKGMIV 599

QY 743 TNPGRKSSVRIDQLDREQFNPDIPTPIIVHFGIRPAQLSVAGDPQYOKLWKSYYKLKH 802
 600 TNPGRKSSVRIDQLDREQFNPDIPTPIIVHFGIRPAQLSVAGDPQYOKLWKSYYKLKH 659

QY 803 LLANSKXVOTDKQKLAOREEALQKTRKQTMREVTVELSSGCFKGTGRSVCOHAMM 862
 660 LLANSKXVOTDKQKLAOREEALQKTRKQTMREVTVELSSGCFKGTGRSVCOHAMM 719

QY 863 LPLVLTNHRHOCIMLIDKLGTVFODRCLQLAMTHPSHNLFGMNPDHARSLSNCGI 922
 720 LPLVLTNHRHOCIMLIDKLGTVFODRCLQLAMTHPSHNLFGMNPDHARSLSNCGI 779

QY 923 ROPKYGDRKVNHHMKKGINVLINIMSRIGODPTPSRIINHERLEPLGDAVVEFLTSV 982
 780 ROPKYGDRKVNHHMKKGINVLINIMSRIGODPTPSRIINHERLEPLGDAVVEFLTSV 839

QY 983 HLYULFESLBEGLAYRTAIVONOHMLAMAKKLEIDRFLVAHGPDLCESDLRIAMAN 1042
 840 HLYULFESLBEGLAYRTAIVONOHMLAMAKKLEIDRFLVAHGPDLCESDLRIAMAN 899

QY 1043 CFEALIGAVYLEGSLBEAKQLFGRLTFNDPDLREVWLYNPLHPLQOEPTDROLIETSP 1102

DB 900 CFEALIGAVYLEGSLBEAKQLFGRLTFNDPDLREVWLYNPLHPLQOEPTDROLIETSP 959
 QY 1103 VLOKLTPEEALGVITHTHYRLIARAFTLRTVGFNNHILGHNOEMELGDSIMOLVATEYL 1162
 DB 960 VLOKLTPEEALGVITHTHYRLIARAFTLRTVGFNNHILGHNOEMELGDSIMOLVATEYL 1019

QY 1163 FIHFPHHGHULTLRSSLVNNRTQAKVAEELQMOEYATINDTKRPVGLRTKTLADLLE 1222
 1020 FIHFPHHGHULTLRSSLVNNRTQAKVAEELQMOEYATINDTKRPVGLRTKTLADLLE 1079

QY 1223 SFIAALYTDKLELYVTFEYVAVCFPPRLKEPILNQDNDPKSQLOCCCLTRTEGKEDIP 1282
 1080 SFIAALYTDKLELYVTFEYVAVCFPPRLKEPILNQDNDPKSQLOCCCLTRTEGKEDIP 1139

QY 1283 LYKTLQTVGSPSHARTVAVYFKEGRIQCGKGPSIQOAEWGAAMDLEKYNFQMAHOKR 1342
 1140 LYKTLQTVGSPSHARTVAVYFKEGRIQCGKGPSIQOAEWGAAMDLEKYNFQMAHOKR 1199

QY 1343 FIGRKYRQELKEMWREHOREPDEDEDIKK 1374
 DB 1200 FIERKTYRQELKEMWREHOREPDEDEDIKK 1231

RESULT 4
 Q5ZIR3 CHICK PRELIMINARY; PRT; 1336 AA.
 ID Q5ZIR3 CHICK PRELIMINARY; PRT; 1336 AA.
 AC Q5ZIR3
 DT 25-OCT-2004 (TEMBLrel. 28, Created)
 DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=RCJMB04.24a21;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxId=9031;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CB; TISSUE=burba;
 RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
 RA Fiedler P., Kutter S., Blagoderaki A., Kostovska D., Kotter M.,
 RA Plachy J., Carinci P., Hayashizaki Y., Buerstedde J.M.;
 RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
 RT gene function analysis";
 RL Genome Biol. 6:R6-R6(2005).
 DR EMBL; AJ720721; CAG32380.1; -; mRNA.
 DR GO; GO:0005634; Cnucleus; IEA.
 DR GO; GO:0003725; P.double-stranded RNA binding; IEA.
 DR GO; GO:0004519; Pendonuclease activity; IEA.
 DR GO; GO:0016787; F.hydrolase activity; IEA.
 DR GO; GO:0004525; F.ribonuclease III activity; IEA.
 DR GO; GO:0006396; P.RNA processing; IEA.
 DR InterPro; IPR001159; DS_RBD.
 DR InterPro; IPR000999; RNase_III.
 DR Pfam; PF00035; dsm; 1.
 DR Pfam; PF00636; Ribonuclease_3; 2.
 DR SMART; SM00358; DSRM; 1.
 DR SMART; SM00535; RIBOC; 2.
 DR PROSITE; PS0137; DS_RBD; 1.
 DR PROSITE; PS00517; RNASE_3_1; 2.
 DR PROSITE; PS0142; RNASE_3_2; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 1336 AA; 154836 MW; 50D9A9A9C53DE0CF CRC64;

Query Match 84.6%; Score 6344.5; DB 2; Length 1336;
 Best Local Similarity 85.3%; Pred. No. 1.7e-301;
 Matches 1171; Conservative 76; Mismatches 81; Indels 45; Gaps 5;

QY 10 MSFHPGCGPRGCGHGAAPSASFRPQVRLHLPQOPVQYQYEPSPASTTFSSPAP 69
 DB 1 MSFHPGCGPRGCGHGAAPSASFRPQVRLHLPQOPVQYQYEPSPASTTFSSPAP 60

[illegible]

Oy	1142	HNOMEEFLGSIIMOLVATEYVIFHPDPHBSHLLTLSSLVNNNTQAKVAEILGMOEYAI	1201
Db	1104	HNOMEEFLGDSIMOLVATEYVIFHPDPHBSHLLTLSSLVNNNTQAKVAEILGMOEYAI	1165
Oy	1202	TNDTKKPVGRTKTYTLTDLLESFTALVYTDKOLEVYHTFMNVCFEPRLKEFTLNDWNP	1261
Db	1164	TNDTKKPVALRTTTLTDLLESFTALVYTDKOLEVYHTFMNVCFEPRLKEFTLNDWNP	1223
Oy	1262	KSQIQOCCCLTTRTEGKEPDIPLXYTLQTVGSHARTTVAVYFGERIGCGKGSIOQAE	1332
Db	1224	KSQIQOCCCLTTRTEGKEPDIPLXYTLQTVGSHARTTVAVYFGERIGCGKGSIOQAE	1283
Oy	1322	MGAMMDALEKNFPQMAHOKRFIRKTYROELKEKRWEREHOERPEDETIKK	1374
Db	1284	MGAMMDALEKNFPQMAHOKRFIRKTYROELKEKRWEREHOERPEDETIKK	1336
RESULT 5			
O80269_MOUSE PRELIMINARY; PRT; 1021 AA.			
ID	O80269_MOUSE		
AC	O80269		
DT	01-JUN-2003 (TREMBLrel. 24, Created)		
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)		
DE	Ribonuclease III.		
GN	Name=Echb12; Synonyms=Rnaasen;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridea; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=BALB/c; TISSUE=Kidney;		
RX	MEDLINE=22730743; PubMed=12191433;		
RA	Fortin K.R., Nicholson R.H., Nicholson A.W.;		
RT	"Mouse ribonuclease III. CDNA structure, expression analysis, and		
RT	chromosomal location."		
RL	BMC Genomics 3:26-26(2002).		
DR	EMBL; AF533013; AAC65849.1; -. mRNA.		
DR	HSSP; O67082; LUFZ.		
DR	Ensembl; ENSMUSG0000022191; Mus musculus.		
DR	MGI; MGI:1261425; Echb12.		
DR	MGI; MGI:1261425; Rnaasen.		
DR	GO; GO:0005622; C:intracellular; IEA.		
DR	GO; GO:0003725; F:double-stranded RNA binding; IEA.		
DR	GO; GO:0004525; F:ribonuclease III activity; IEA.		
DR	GO; GO:0006356; P:RNA processing; IEA.		
DR	InterPro; IPR001159; De RNA bd.		
DR	InterPro; IPR000999; Rnaase_III.		
DR	Pfam; PF00035; dnm. 1.		
DR	Pfam; PF00636; Ribonuclease_3. 2.		
DR	SMART; SM00358; DSRW; 1.		
DR	SMART; SM00355; RIBOC; 2.		
DR	PROSITE; PS00137; DS_RBD; 1.		
DR	PROSITE; PS00517; RNAS3_1; 2.		
DR	PROSITE; PS00142; RNASE_3_2; 2.		
DR	SEQUENCE 1021 AA; 118200 MW; 15E4899270AF6F29 CRC64;		
Query Match 69.6%; Score 5217; DB 2; Length 1021;			
Best Local Similarity 73.1%; Pred. No. 1.5e-246;			
Matches 1003; Conservative 10; Mismatches 8; Indels 352; Gaps 1			
Oy	2	MOGNTCHMSHPRGRCGRGCGHARGASAPRPONRLTHPOOPVOYOEPPSARST	61
Db	1	MOGNTCHMSYHPRGCGRCGRGCGHARGASAPRFRONRLTHPOOPVOYOT	52
Oy	62	TFSNSPAPNPLPRPBDVFVFPFPMPSAQGLPCPIRPPFNQMRAPFVYPCFPMP	121
Db	53	-----	52
Oy	122	PMPCPNPNPVGAPPGCGTFPFMMPPSPMHPPPPPVMDQVNTQYPPGSHNFP	181

Db	53	-----	52	-----	52
Qy	182	FNSFONNSSFLPSANNSSPHFRLLPYPYLPKASERSRPRLKGYDHHRHDSHRC	241	-----	52
Db	53	-----	52	-----	52
Qy	242	ERHRSGLDRERGRSPDRRDRGRYRSDYDRGRTPSRHRSYERSRERERHRRHNRASP	301	-----	52
Db	53	-----	52	-----	52
Qy	302	SLERSYKKEKXKSGRSYGLSVPEBAGCTPBLPGELIKNTDSWAPPLEIVNHRSPSREK	361	-----	52
Db	53	-----	52	-----	52
Qy	362	RARWEEKDRSDNOSGCKDKNTYSIKEKEPETMPDKNEBEELLRPMVIRCHSANY	421	-----	52
Db	53	-----	68	-----	52
Qy	422	YSSDPMDOVGSTVYVGRSLRDLVYKFEELIGROEKAKAARPPPEPKTKLDEJLESS	481	-----	52
Db	69	YSSDPMDOVGSTVYVIGLNDLVYKFEELIGNRKAKAARPPPEPKTKLDEJLESS	128	-----	52
Qy	482	ESECESDSTCSSSSSDSEVPDVIAEIRKKKAPRDLDELMYNDPGOMDGPLCKCSAK	541	-----	52
Db	129	ESECTDDSTCSSSSDSEVFVIAEIRKKKAPRDLDELMYNDPGOMDGPLCKCSAK	188	-----	52
Qy	542	ARTRGIRHSIYGEBAIKPCRMTNNAARLPHYRITVSPPTNFLTRDPTVIEYDDHEIF	601	-----	52
Db	189	ARTRGIRHSIYGEBAIKPCRMTNNAARLPHYRITVSPPTNFLTRDPTVIEYDDHEIF	248	-----	52
Qy	602	EGSFMAFAPLTTNPLCKVIRPNIDYTHFIEEMMPENPCVKGLELSPLPRODLEIYD	661	-----	52
Db	249	EGSFMAFAPLTTNPLCKVIRPNIDYTHFIEEMMPENPCVKGLELSPLPRODLEIYD	308	-----	52
Qy	662	WNLKGPLFEDSPCCPRFHFMPRFVRLPDGKEVLSMHQILLYLRCSKALVPEEBIAN	721	-----	52
Db	309	WNLKGPLFEDSPCCPRFHFMPRFVRLPDGKEVLSMHQILLYLRCSKALVPEEBIAN	368	-----	52
Qy	722	MLQWELEWQKVAEECKGMIVTNPCTKSSVARIQDLREQFNPVITPEIIVHFGIRPAQ	781	-----	52
Db	369	MLQWELEWQKVAEECKGMIVTNPCTKSSVARIQDLREQFNPVITPEIIVHFGIRPAQ	428	-----	52
Qy	782	LSYAGDPQOYLKMSYVVLRLHLANSPRVKQIQDQKLAQREBALOKIQKQTMREXTVE	841	-----	52
Db	429	LSYAGDPQOYLKMSYVVLRLHLANSPRVKQIQDQKLAQREBALOKIQKQTMREXTVE	488	-----	52
Qy	842	LSOGFMTGTGRSDVCOHAMMLPVLTNHIRYKQCLMHLDKLIGYFEOORCLLOLANTPS	901	-----	52
Db	489	LSOGFMTGTGRSDVCOHAMMLPVLTNHIRYKQCLMHLDKLIGYFEOORCLLOLANTPS	548	-----	52
Qy	902	HHLNFGNPDRHARNSLSNCGIRQPKYGRKYNHMMRKKGINTLINISRLGODPTPSR	961	-----	52
Db	549	HHLNFGNPDRHARNSLSNCGIRQPKYGRKYNHMMRKKGINTLINISRLGODPTPSR	608	-----	52
Qy	962	INNHRELFEGDAVVEFLTSVHLTYLPSLEEGGLATYRTIAIVONQHILMLAKKLELDPF	1021	-----	52
Db	609	INNHRELFEGDAVVEFLTSVHLTYLPSLEEGGLATYRTIAIVONQHILMLAKKLELDPF	668	-----	52
Qy	1022	MLVYHGPRLCGESDLRHMANCFEALIAVYVLESLLEAKOLFGLLFPNDDBLEVMYNY	1081	-----	52
Db	669	MLVYHGPRLCGESDLRHMANCFEALIAVYVLESLLEAKOLFGLLFPNDDBLEVMYNY	728	-----	52
Qy	1082	PLHPLQLOEPMTDRQILETSPLQLCTFEFEALIGVYFNHVLLARAFLRTVGFNHLTLG	1141	-----	52
Db	729	PLHPLQLOEPMTDRQILETSPLQLCTFEFEALIGVYFNHVLLARAFLRTVGFNHLTLG	788	-----	52
Qy	1142	HNQRMFLGDSIMOLVATEYLFIHFRDHHBGLTLRSLVNNRTQAVAEELGQOEYAI	1201	-----	52
Db	789	HNQRMFLGDSIMOLVATEYLFIHFRDHHBGLTLRSLVNNRTQAVAEELGQOEYAV	848	-----	52
Qy	1202	TNDTKTRVGLRTKTLADLESFIALYTDQDLVYVHFMVNCPPPRKESITLQDMMDP	1261	-----	52
Db	849	TNDTKTRVGLRTKTLADLESFIALYTDQDLVYVHFMVNCPPPRKESITLQDMMDP	908	-----	52

Qy	1262	ISOLQOCCILTRTGEKEDPILPYKTLQTVGSHARITVAVYFPGERGCGCKGSIQOAE	13274
Db	909	KSQIQOCCILTRTGEKEDPILPYKTLQTVGSHARITVAVYFPGERGCGCKGSIQOAE	968
Qy	1322	MGAAMDALKEKNFPQMAHQKRFIFGRKYRQELKEMRWEREHQERPEDEDIDIK	1374
Db	969	MGAAMDALKEKNFPQMAHQKRFIFGRKYRQELKEMRWEREHQERPEDEDIDIK	1021
RESULT 6			
Q6PF88	MOUSE	PRELIMINARY;	PRT; 942 AA.
AC	Q6PF88		
DT	05-JUL-2004	(TREMBLrel. 27, Created)	
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)	
DE	Rnaasen protein (fragment).		
GN	Name=Etch12; Synonyms=Rnaasen;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;		
OC	Muridea; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Czech II:		
RC	TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;		
RX	MEDLINE=22388325; PubMed=12477933; DOI=10.1073/pnas.942603899;		
RA	Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton W., Soares M.B., Bonaldo M.F., Casavant J.L., Scheetz T.E.,		
RA	Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Rata S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Boeak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield J.S.N., Krzywinski M.I., Skelton A., Smallus D.E.,		
RA	Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Czech II:		
RC	TISSUE=Mammary tumor metastatized to lung. Tumor arose spontaneously;		
RG	NIH WGC Project;		
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC057687; AAHS7687.1; -, mRNA.		
DR	Ensembl; ENSMUSG00000022191; Mus musculus.		
DR	MGI; MGI:1261425; Etch12.		
DR	MGI; MGI:1261425; Rnaasen.		
DR	GO; GO:0005622; C:intracellular; IEA.		
DR	GO; GO:0003755; F:double-stranded RNA binding; IEA.		
DR	GO; GO:0004525; F:ribonuclease III activity; IEA.		
DR	GO; GO:0006396; P:RNA processing; IEA.		
DR	InterPro; IPR001159; ds_RNA_bd.		
DR	InterPro; IPR000999; RNase_III.		
DR	Pfam; PF00035; dsrm_1.		
DR	Pfam; PF00636; Ribonuclease_3; 2.		
DR	SMART; SM00358; DSRM; 1.		
DR	SMART; SM00535; RIBOC; 2.		
DR	PROSITE; PS01317; DS_RBD; 1.		
DR	PROSITE; PS00517; RNase_3_1; 2.		
DR	PROSITE; PS0142; RNase_3_2; 2.		
FT	NON_TER 1		

SQ SEQUENCE 942 AA; 109445 MW; C3629E4676FC8CB CRC64;
 Query Match 66.4%; Score 4982; DB 2; Length 942;
 Best Local Similarity 98.9%; Pred. No. 4,1e-235;
 Matches 932; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 433 GTVVGTSRLDLYDKFEELSGROEKAKARPPWEPKTKLDBDLSSESCEDESDST 492
 DB 1 STVVGTSRLDLYDKFEELSGROEKAKARPPWEPKTKLDBDLSSESCEDESDST 60
 QY 493 CSSSSSEVFVDAIEIKKKKAPDRDLDELWMDPQNMNDGPICKCSAKARRTGIRHSY 552
 DB 61 CSSSSSEVFVDAIEIKKKKAPDRDLDELWMDPQNMNDGPICKCSAKARRTGIRHSY 120
 QY 553 PEEBAIKPCRPMTNAGRLFHYRITYSPPTNPLTDPTVIEYDDHEIIEGFSMEFHAHL 612
 DB 121 PEEBAIKPCRPMTNAGRLFHYRITYSPPTNPLTDPTVIEYDDHEIIEGFSMEFHAHL 180
 QY 613 TNIPICKVIRFNIDYTHIFIEEMMPENFCVKGLESLFLPNDILELYMNLKGPLFEDS 672
 DB 181 TNIPICKVIRFNIDYTHIFIEEMMPENFCVKGLESLFLPNDILELYMNLKGPLFEDS 240
 QY 673 PCCCRPFHMPFRVRLPDGKKEVLSMHQILYLKCSKALVPEEIRIAMLQMEELMOK 732
 DB 241 PCCCRPFHMPFRVRLPDGKKEVLSMHQILYLKCSKALVPEEIRIAMLQMEELMOK 300
 QY 733 YAECKGMIVTNPGRSPSVRIDQDREDFNPVTFPIIVHFGIRPAQLSYAGDPQYOK 792
 DB 301 YAECKGMIVTNPGRSPSVRIDQDREDFNPVTFPIIVHFGIRPAQLSYAGDPQYOK 360
 QY 793 LMKSYVKLHLLANSKPKVQTDQKLAOREBALQTRQKXWREYTVELSSCGFWKTGI 852
 DB 361 LMKSYVKLHLLANSKPKVQTDQKLAOREBALQTRQKXWREYTVELSSCGFWKTGI 420
 QY 853 RSDVCOHAMMLPVLTTHIRYHOCIAMLDKLGTFPDRCLLOIAMHPSHHLNFGNPPH 912
 DB 421 RSDVCOHAMMLPVLTTHIRYHOCIAMLDKLGTFPDRCLLOIAMHPSHHLNFGNPPH 480
 QY 913 AANSLNSCGIRQPKYGDRAKVNHHMKKGINTLINIMSRIGODDPTPSRINHERLEFG 972
 DB 481 AANSLNSCGIRQPKYGDRAKVNHHMKKGINTLINIMSRIGODDPTPSRINHERLEFG 540
 QY 973 DAVEEFLTSVHLYYLPPSLBEGGLATYRTAIVONQHILAMAKKLEIDRPMLYAHGPDLCR 1032
 DB 541 DAVEEFLTSVHLYYLPPSLBEGGLATYRTAIVONQHILAMAKKLEIDRPMLYAHGPDLCR 600
 QY 1033 EBDLRAMANCERALLGAVYLESGLEBAKQLEGRLLFNDPDLREVLANTPLHPLQOEEN 1092
 DB 601 EBDLRAMANCERALLGAVYLESGLEBAKQLEGRLLFNDPDLREVLANTPLHPLQOEEN 660
 QY 1093 TDRQLIETSPVLOKLEFEBAIGVIFTHYRLARATLRTVGFNHLTLGHNMREFLAGS 1152
 DB 661 TDRQLIETSPVLOKLEFEBAIGVIFTHYRLARATLRTVGFNHLTLGHNMREFLAGS 720
 QY 1153 IMQLVATEYLFIHPDHEGHLTLRSSLVNNFTQAKVAEELGMOEYAITNDXKRPVGL 1212
 DB 721 IMQLVATEYLFIHPDHEGHLTLRSSLVNNFTQAKVAEELGMOEYAITNDXKRPVGL 780
 QY 1213 RRTKLADLESFALAYTDKLELYVTHFMNVCFFPLKKEITLNQNMNDPSQIQCCCLTL 1272
 DB 781 RRTKLADLESFALAYTDKLELYVTHFMNVCFFPLKKEITLNQNMNDPSQIQCCCLTL 840
 QY 1273 RTEGKRPDIPLYKTLQTVGSHARTYTVAVYFKGEIRIGCGKSPISQOAEKGAMADLEKY 1332
 DB 841 RTEGKRPDIPLYKTLQTVGSHARTYTVAVYFKGEIRIGCGKSPISQOAEKGAMADLEKY 900
 QY 1333 NFPQMAHQRFIRGRKYRQELKEMRWERHQERPEDETDIKK 1374
 DB 901 NFPQMAHQRFIRGRKYRQELKEMRWERHQERPEDETDIKK 942
 RESULT 7
 07TM18_MOUSE

ID 07TM18_MOUSE PRELIMINARY; PRT; 780 AA.
 AC 07TM18;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Naaen protein (Fragment).
 GN Name=Bchl2; Synonym=Naaen;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Mouse;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haich F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stachenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Mullahy S.J.,
 RA Baha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rubin G.M., Rodriguez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Roderfeldt Y.S.N., Krzywicki M.T., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Mouse;
 RG NIH MGC Project;
 RL Submitted (Aug-2003) to the EMBL/Genbank/DBA databases.
 DR EMBL; BC055696; AAH55696.1; -; mRNA.
 DR HSPF; 067082; 1UTZ.
 DR Ensembl; ENSMUSG0000022191; Mus musculus.
 DR MGI; MGI:1261425; Etoh12.
 DR MGI; MGI:1261425; Naaen.
 DR GO; GO:0006222; C:intracellular; IEA.
 DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
 DR GO; GO:0004525; F:ribonuclease III activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR InterPro; IPR001159; Ds_RNA_Bd.
 DR InterPro; IPR000999; Naaen_III.
 DR Pfam; PF00035; dsrm; 1.
 DR Pfam; PF00636; Ribonuclease_3; 2.
 DR SMART; SM00358; DSRM; 1.
 DR SMART; SM00355; RIBOC; 2.
 DR PROSITE; PS00137; Ds_RBD; 1.
 DR PROSITE; PS00517; RNASE_3_1; 2.
 DR PROSITE; PS00142; RNASE_3_2; 2.
 FT NON TER 1
 SQ SEQUENCE 780 AA; 91045 MW; 78D734332C75382E CRC64;
 Query Match 54.9%; Score 421; DB 2; Length 780;
 Best Local Similarity 99.1%; Pred. No. 3.8e-193;
 Matches 773; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 595 DDEHYTFEGFSFPAHAPLTNIPICKVIRFNIDYTHIFIEEMMPENFCVKGLESLFLR 654
 DB 1 DDEHYTFEGFSFPAHAPLTNIPICKVIRFNIDYTHIFIEEMMPENFCVKGLESLFLR 60
 QY 655 DILELYDMNLKGPLFEDSPCCPRFHFMPFRVRLPDGKKEVLSMHQILYLKCSKALV 714
 DB 61 DILELYDMNLKGPLFEDSPCCPRFHFMPFRVRLPDGKKEVLSMHQILYLKCSKALV 120

```
QY 715 PEEBIAANMLQWELEMOXKAEBCGMIVTNPGTKPSSVRIDQDREQFNPDIPTPIIYN 774
|||
DB 121 PEEBIAANMLQWELEMOXKAEBCGMIVTNPGTKPSSVRIDQDREQFNPDIPTPIIYN 180
QY 775 FGRPAQULSYAGDPOYOKLMSYVYKRLHLLANSPPKQDQKQDLAOREEALQKIRQKNTM 834
|||
DB 181 FGRPAQULSYAGDPOYOKLMSYVYKRLHLLANSPPKQDQKQDLAOREEALQKIRQKNTM 240
QY 835 RREVTVELSSQGFWMKGISDVQCHAMMLPVLTHTIRYQCLMHLQKLGTYQDQCLQ 894
|||
DB 241 RREVTVELSSQGFWMKGISDVQCHAMMLPVLTHTIRYQCLMHLQKLGTYQDQCLQ 300
QY 895 LAMTHPSHHLNFGMNDHANSLNSGCIROPKYDKRKVHMHMKXGINTLJINMSRLQ 954
|||
DB 301 LAMTHPSHHLNFGMNDHANSLNSGCIROPKYDKRKVHMHMKXGINTLJINMSRLQ 360
QY 955 DDPTRIRINHNLELELGDVAVVEFLTSVHLXYLPESLEBGLATYTTAIVONHMLAK 1014
|||
DB 361 DDPTRIRINHNLELELGDVAVVEFLTSVHLXYLPESLEBGLATYTTAIVONHMLAK 420
QY 1015 KLELDPMFLYANGPDLGRESDDLHMANCPREALIGAVLEGSLEBAKOLFGRLLFNDPDL 1074
|||
DB 421 KLELDPMFLYANGPDLGRESDDLHMANCPREALIGAVLEGSLEBAKOLFGRLLFNDPDL 480
QY 1075 REVWLNYPRLPQLQBPNTDROLIETSPVLOKLTREBEAIGVIFTHVRLARAFTLRTVG 1134
|||
DB 481 REVWLNYPRLPQLQBPNTDROLIETSPVLOKLTREBEAIGVIFTHVRLARAFTLRTVG 540
QY 1135 FNNHLLTGNHQRMEFLDSTIMQVATETYLTHPPDHHEGHLTLRSSLVNNRTOAKYAEEL 1194
|||
DB 541 FNNHLLTGNHQRMEFLDSTIMQVATETYLTHPPDHHEGHLTLRSSLVNNRTOAKYAEEL 600
QY 1195 GNOEYAITNDKTKRPVGLRTKTLADLESFIALYTXDLEVYHTPMVCFPRLEKFTL 1254
|||
DB 601 GNOEYAITNDKTKRPVGLRTKTLADLESFIALYTXDLEVYHTPMVCFPRLEKFTL 660
QY 1255 NODMNDPKSOLOOCCLTLRTEGKEPDIPLYKTLQTPSGHARTYVAVYFKGERLGGKG 1314
|||
DB 661 NODMNDPKSOLOOCCLTLRTEGKEPDIPLYKTLQTPSGHARTYVAVYFKGERLGGKG 720
QY 1315 PSIOQAEMGAMDALEKYNPQMAHQKRIKRYROELKEMWEREHOBERDETDIKK 1374
|||
DB 721 PSIOQAEMGAMDALEKYNPQMAHQKRIKRYROELKEMWEREHOBERDETDIKK 780

RESULT 8
QASKW1_TETNG PRELIMINARY; PRT, 1041 AA.
ID QASKW1_TETNG
AC QASKW1;
DT 13-SEP-2005 (TREMBLREL 31, Created)
DT 13-SEP-2005 (TREMBLREL 31, Last sequence update)
DE 13-SEP-2005 (TREMBLREL 31, Last annotation update)
ORFNames=GSTENG0016530001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallouf O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Maucell E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaife D., Flaher S., Lutfalla G., Dossat C., Segutens B.,
RA Daesila C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Antouane V., Jubin C., Castell V., Katinka M., Vachele B.,
RA Blemont C., Steill Z., Catolico L., Poulain J., Berardins V.,
RA Craud C., Duprat S., Broctier P., Coucanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappe C., McKernan K.J., McGowan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Nestrov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
```

```
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RC Submitted (FEB-2004) to the EMBL/GenBank/DBJ database.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEB014565; CAF98721.1; -; Genomic DNA.
SQ SEQUENCE 1041 AA; 120534 MW; 2CA7A165A7BB3CDB CRC64;

Query Match 50.2%; Score 3767; DB 2; Length 1041;
Best Local Similarity 67.5%; Pred. No. 1,1e-175;
Matches 728; Conservative 80; Mismatches 135; Indels 140; Gaps 18;

QY 10 MSFHPGCGPRGGHGAAPSAPSPQNLRLHPQOPVOYQYEPSPASTTF-SNSPA 68
|||
DB 1 MSFHPDRAAHRGMAAGPHMSPOHYRPTLRLRPPRPVRYHYDQVPPSSGYHNS-- 58
QY 69 PNFLLPRDPFVPPPP-----PMPSAQGP-LPPCPILRPPRNQMHNFVPPPC 116
|||
DB 59 -GYMHPSTFMQYSPGAFADSRSEPVARQEGPRLPCLPTPTSLYMMPTVPPRPLA 117
|||
QY 117 PPMPPMPMCPN---NPVPGAPRQCGTFRPMMPMSMHPPPPPVMPQGVVYQYRYS 173
|||
DB 118 PPMPPAPLHSPKSTHPPY-----PMNSHPHPPP----- 147
QY 174 HNNFPSPFNS-FQNNPSSFLPSANSSPHRHLPVYPLPYAPSERSPERLKHYYDHR 232
|||
DB 148 ---FPPPTNSGHHVQSGSFADLG-----FRHGGSPYGHDXPADXRGRPREGYHDDL 198
QY 233 HEDHSH-GEGERHSL---DREERGRSPDR--RQDSRYRSDYDRGRTSRHSYKRSR 286
|||
DB 199 QGYSHSGHSDQKMEFFDGRDGRSPDRRRPREGGRHSEYDRKRTPPRH---RSRE 254
|||
QY 287 RREBRHHRDNRSPSLERSYKKEYRSGRSGVLSVPEPAGCTPRLPEIINKTDSWAP 346
|||
DB 255 RSRERFRHDSRSOSPDHRRKPRSRSG----- 283
QY 347 PLEIVNHRSPSEKKAARWEERKDRWSDQSS-GKDKNTYSIKEKEPTEMDK----- 399
|||
DB 284 -----SRERGRSRWEERERRESSESSAPSRGRSYVSANRSEAVREREGCRP 333
|||
QY 400 -----NEEBEELLKPVWIRCTHSENYSSDPMDQVGDSTVGT 438
|||
DB 334 TGRGEGPALPEHDEKEKEGEKEEELLKPAWIRCTHAESYSSNDPMDQVGDSTVGT 393
QY 439 SLRLDYKFEFEEELSGROKAKAARPPWPPTKIDEDL-ESSSEBECSSDED-STCSS 496
|||
DB 394 SKLRDLYORFEELGRRQARAARPKWPTKIDEDDESSSEBECSSDADGSCSS 453
QY 497 SDSEYDVYAEIKRKAHPDRLLHDELYNDPCQMDGPKCSAKXARRGIRHSIYGSSE 556
|||
DB 454 SSDVDFVDAEIKRKAHNDRLHEELLYNDPCQMDGPKCSAKXARRGIRHSIYGSSE 513
QY 557 AIKPCRPMTNNAQRLFHYRITVSPPTNFLTDRSPVIEYDHEYIFEGFSMAAPLTNP 616
|||
DB 514 SVKQCRAMNNAGKLFHYRITVSPPTNFLTDRSPVIEYDHEYIFEGFSHTPLTNIP 573
QY 617 LCKVIRENDYTHIFEEEMBPENFCYKGLFSLFRLILLYMNLGRLPEEDSPCC 676
|||
DB 574 LCRVIRENDYTHIFEEEMBPENFCYKGLFSLFRLILLYMNLGRLPEEDSPCC 633
QY 677 PRFHMPFRVRLPDGKGVLSMHOILYLRLCSKALVBEERIANMLQWELEMOXKAE 736
|||
DB 634 PQHMPFRVRLPDGKGVLSMHOILYLRLCSKALVBEERIANMLQWELEMOXKAE 693
QY 737 CGKMITVNPGRTPSSVRIDQDREQFNPDIPTPIIYNFGRPAQULSYAGDPOYOKLMS 796
|||
```

Db 694 CKGMIVTNGMKRPSVRIDQLDRQFNPVITFPPIIVHFGIRPAQLSYAGDPOYOKLWKS 753
QY 797 YKCLRHLLANSPKVKQTDKOKLAOREALOKIROKNTMRREVVELSSGFWKTIQRSDY 856
Db 754 YKCLRHLLANSPKVKQTDKOKLAOREALOKIROKNTMRREVVELSSGFWKTIQRSDY 813
QY 857 COHAMMLPVLTHIRHYOCLMHLDKLIGYTFQDRCLQLQAMTHPSHHLNFGNPDHANS 916
Db 814 COHAMMLPVLTHIRHYOCLMHLDKLIGYTFQDRCLQLQAMTHPSHHLNFGNPDHANS 873
QY 917 LSNCGIRORPKYGRKRVHNMHMKKGINTLINIMSLRGODDPSPSRINNHRELFEGDAVY 976
Db 874 LSNCGIRORPKYGRKRVHNMHMKKGINTLINIMSLRGODDPSPSRINNHRELFEGDAVY 933
QY 977 EFLTSVHLVYLFPSLEEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGPDLCRESDL 1036
Db 934 EFLTSVHLVYLFPSLEEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGPDLCRESDL 993
QY 1037 RHMANCFEALIG 1049
Db 994 RHMANCFEALIG 1006

RESULT 9
Q80U5 MOUSE PRELIMINARY; PRT; 705 AA.
ID Q80U5.1
AC 080U5.1
DT 01-JUN-2003 (TREMBlrel. 24, Last Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Rnaase protein (Fragment).
GN Name=Ech12; Synonyms=Rnaase;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
[1]
[2]
NM NUCLEOTIDE SEQUENCE.
RP STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Caesavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ustin T.B., Tobinlyuk S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakeley R.W., Tuchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
NM NUCLEOTIDE SEQUENCE.
RP STRAIN=129/Sv x 129/Sv-CP; TISSUE=ES Cell;
RC NIH MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC050057; AAH50057.1; -; mRNA.
DR HSSP; 067082; 1JFZ.
DR Ensembl; ENSMUSG00000022191; Mus musculus.
DR MGI; MGI:1261425; Etch12.
DR MGI; MGI:1261425; Rnaase.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.

DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; Ds_RNA_bd.
DR InterPro; IPR000999; Rnaase_III.
DR Pfam; PF00035; darm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; Rnaase_3_1; 2.
DR PROSITE; PS50142; Rnaase_3_2; 2.
FT NON_TER
SQ SEQUENCE 705 AA; 82084 MW; 3332C69A5F1977D6 CRC64;
Query Match 49.3%; Score 3696; DB 2; Length 705;
Best Local Similarity 98.7%; Pred. No. 2e-172;
Matches 696; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 670 EDSPPCCRRFHFMPFVFLPDGKREYVSMQIILYLRCSKALVPEEIANMLQWEELE 729
Db 1 EDSPPCCRRFHFMPFVFLPDGKREYVSMQIILYLRCSKALVPEEIANMLQWEELE 60
QY WOKYAECKGMIVTNGMKRPSVRIDQLDRQFNPVITFPPIIVHFGIRPAQLSYAGDPQ 789
Db WOKYAECKGMIVTNGMKRPSVRIDQLDRQFNPVITFPPIIVHFGIRPAQLSYAGDPQ 120
QY 790 YOKLWKS YVYKLRHLLANSPKVKQTDKOKLAOREALOKIROKNTMRREVVELSSGFWK 849
Db WOKYAECKGMIVTNGMKRPSVRIDQLDRQFNPVITFPPIIVHFGIRPAQLSYAGDPQ 120
QY 121 YOKLWKS YVYKLRHLLANSPKVKQTDKOKLAOREALOKIROKNTMRREVVELSSGFWK 180
QY 850 TGISRDVCOHAMMLPVLTHIRHYOCLMHLDKLIGYTFQDRCLQLQAMTHPSHHLNFGN 909
Db TGISRDVCOHAMMLPVLTHIRHYOCLMHLDKLIGYTFQDRCLQLQAMTHPSHHLNFGN 240
QY 181 TGISRDVCOHAMMLPVLTHIRHYOCLMHLDKLIGYTFQDRCLQLQAMTHPSHHLNFGN 240
QY 910 PDHANSLSNCGIRORPKYGRKRVHNMHMKKGINTLINIMSLRGODDPSPSRINNHRELE 969
Db PDHANSLSNCGIRORPKYGRKRVHNMHMKKGINTLINIMSLRGODDPSPSRINNHRELE 300
QY 241 PDHANSLSNCGIRORPKYGRKRVHNMHMKKGINTLINIMSLRGODDPSPSRINNHRELE 300
QY 970 FLGDAVVEFLTSVHLVYLFPSLEEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGPD 1029
Db FLGDAVVEFLTSVHLVYLFPSLEEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGPD 360
QY 301 FLGDAVVEFLTSVHLVYLFPSLEEGGLATYRTAIVONOHMLAKKLELDPEMLYAHGPD 360
QY 1030 LCRESDDLPHANANCFEALIGAVYLEGSLBEAKQLFGILLPNDPDLREVWLANPLHPLOQ 1089
Db LCRESDDLPHANANCFEALIGAVYLEGSLBEAKQLFGILLPNDPDLREVWLANPLHPLOQ 420
QY 361 LCRESDDLPHANANCFEALIGAVYLEGSLBEAKQLFGILLPNDPDLREVWLANPLHPLOQ 420
QY 1090 EENTROLIETSPVLQKLTPEFEALIGVFTYHRLARAFLLRTVGFNHLTLGHQRMETL 1149
Db EENTROLIETSPVLQKLTPEFEALIGVFTYHRLARAFLLRTVGFNHLTLGHQRMETL 480
QY 421 EENTROLIETSPVLQKLTPEFEALIGVFTYHRLARAFLLRTVGFNHLTLGHQRMETL 480
QY 1150 GDSINQIVATEVYLFHPDHEGHLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTKRP 1209
Db GDSINQIVATEVYLFHPDHEGHLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTKRP 540
QY 1210 VGLRKTTLADLIESITIALYTDKLEFYHTFNANVCFPRPKAFINQDNNDKSQLQOCC 1269
Db VGLRKTTLADLIESITIALYTDKLEFYHTFNANVCFPRPKAFINQDNNDKSQLQOCC 600
QY 541 VALRKTTLADLIESITIALYTDKLEFYHTFNANVCFPRPKAFINQDNNDKSQLQOCC 600
QY 1270 LTLRTGKEPDIPLKTTQTVGSPHARTYTVVYKGEISGGKSPSIQAEAGAMDAL 1329
Db LTLRTGKEPDIPLKTTQTVGSPHARTYTVVYKGEISGGKSPSIQAEAGAMDAL 660
QY 1330 EKYNFPOMAHQGRFYGRKYROELKEMWEREHOEREPDETIDIKK 1374
Db EKYNFPOMAHQGRFYGRKYROELKEMWEREHOEREPDETIDIKK 705
RESULT 10
Q9XYN5 DROME PRELIMINARY; PRT; 1327 AA.
ID Q9XYN5.1
AC Q9XYN5.1
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)


```

Db 47 SSKSLDYVYPETPAYASS---VPSYDPYQOPPAYEGEVAYNEQAQKYGQESHYQYQ 103
QY 180 PSFNSFONNPSSEFLPSANNSSSPHRLPPYPL----PKAPSEERSPERLKAYDDHRD 235
Db 104 PA-----SGSFLYES-----XKPPRYPAYSSNRPSEKQ-----RTS 139
QY 236 HSHGGERHRSLDRBERGSPDRRQDSRYRSDYDRGRTSPHRSYERSRERERHRHR 295
Db 140 NSSSGYHHYF-----GYSSGR-----RYEQRHQO-----EHRQLODSRYAHEPRHGAY 183
QY 266 DNRBSPSLERSYKKEYSKSGSYGSLVVEBPAGCTPELBCEIITKPTDSAPPLLEYNRS 355
Db 184 AHRQAKGSGOHY---YGSAAFN-----QVSDVSPGGHHEREN 219
QY 356 PEREKRA--RWEEEKDR-----MSDNQSGDKNYTSIKEKEPEETMPDKNEEELIK 409
Db 220 ETEKTRAPKVEYTERDRLRLRQWCSNFC-----EKPEDEVKKNMLSEDAV 267
QY 410 PWMIRCTHSENYSSDPMDQVDSVTVGTSLRLDYKFEELSGROEKAKARPEWEP 469
Db 268 ESWAVSSPAPLEYERTK-----SENEVRGARLOKCTLPDEELLQRAKVRKELPYVVP 323
QY 470 KTKL-----DEDESSSESCESEDSTCSSSDSEVPDV-----IAEIKRKXKH 514
Db 324 PKKARRRYCKKHKKEBACSSSSSDDDSDDA-----FKLEQDCMBELSKYQH 373
QY 515 PRLDELAYNDPGONNDGPLCKCSAKARTGIRHSIYPGEBAIKPCRPMTNAGRLPAY 574
Db 374 PQRVHADLHNDAGENNDGPLCRCSAKSRIGIRHGIYPGEGYKLCDSNNAAGLFLY 433
QY 575 RTVSPTNFTLDRPTVIEYDDHEYLIEGFSMFAAPLNTNPLCKYIRNIDYTHFIEE 634
Db 434 RISISPTNFTLKTPTIIEKHDEHEFLFEGFSLSHVRLSDLPCKVIRINIEVTEYBE 493
QY 635 MPENPCVKGLEFSFLPRODILELYDNMLKGP-----FEDSPCCRRPHMPPRFRP 690
Db 494 KMPENFTIHELDIIEFKYLFHELELVDFNLMPRLPSGNVEE---CPAHFFPRFRDP 550
QY 691 DQGEKYLSHQOILLYLRCSKALVPEEELANMLQMELEMOKAYEECKMIVTNPCTKS 750
Db 551 DNGKEVLANVEVRYLLDNSAQLVEQOQLHLAQISQSEMONYVDIKMLVTKPKYKC 610
QY 751 SVRIDQLDREOFNPV-----ITPPIVHFGIRPAQLSYAGDPQYQKMLKSYVKLR 801
Db 611 SLRVQDLDN--NSDLPECVDRGTGSHPAIVHFGIHPQLSYAGPEYQKAMREYVKR 668
QY 802 HLLANSFKYQDQKQKLAOREBALQIKRQKNTREYVYELSSQGWKIGISDVQHAM 861
Db 669 HLMANMSKPSFKDKRKLKEKEQRLQEMRTQGRMKRNIYVAISSEGFRTGIMCDVVQHAM 728
QY 862 ML.PV.LTHNLRHQCMLHKLIGYTFODRCLLOLMTHTPSHHLNFGMNPDRHANSLSNG 921
Db 729 LIPVLGHLRFHSLDLLESISYRKRYLLOLALHTHSYKENVYTNDRHANSITNG 788
QY 922 IQPKYGDGRVHMHNRKKGINTLINIMSRIGODPTPSRINNELEFLGDAVVEFLTS 981
Db 789 IQPEBGDKKIHMYNRKGINLVGIMSRFGKEHTVSNITNERLEFLGDAVVEFLSS 848
QY 982 VHLVYLPESLEBGLATTYTAIVQNOHMLAKKLELDPMLYAHDPDLCRESDLRHMA 1041
Db 849 IHLFFMFPELEBGLATYPAIVQNOHMLAKKLELDPMLYAHOSDLCHELELHMA 908
QY 1042 NCEALIGAVVLEGSLEBAKOLFGRLLF--NDPDLREVMYLPYHPLOLOEPNTDQLIE 1100
Db 909 NCEALMGALLDGGIKVADDEVTTDALFRODEKLSIMKNLPHPLOLOEPDLRSCITS 968
QY 1101 SPYLQKLFEEBAIGYLFTHVRLARAFTLTGVPNHLTIGHNORMEFLDSTIMQVATE 1160
Db 969 YARLKLTLFEDSIGIKFKHILARAFTRDSIGFHLTLGSGNQLREPLGDTVLQICSE 1028
QY 1161 YLFIFHPDHHGGLTLRSLVNNRTOAKYAEELGMOEYAI--TNDTKKRPVGLRTTLAD 1219
Db 1029 YLYRHFPEHHEGHLSTLRSSLVNNRTOAVVCDLGMPCYAVVANPK---ADLKTQDRAD 1084

```

```

QY 1220 LIESFIALYTDKLEYHTNTNVCFFPRLKEFLINQPMNDPKSLOOCCLTLR--EKE 1278
Db 1085 LLEAFGLAYVDKGLLYCEOFCHVCLFPRLQFLINQDMNPKSLQOCCLTLRITMDGE 1144
QY 1279 PDILYKTLQVGPBHATYVAVYFKGERIGCGKSPSIOQAEWGAADALEKYN--FPQ 1336
Db 1145 PDIPYKVEASGPNNTVYKAVYFKSRILATSSGSSIOQAEWGAALENSRLFPQ 1204
QY 1337 MAHQKRFYGRKYRQEL---KEWRWERHOREPEDETDIKK 1374
Db 1205 LDHQKRVIAKSKIKQGTGNELNDSDRQHO-----EKIKR 1239

```

RESULT 11

OSTOKA ANOGA

ID OSTOKA ANOGA PRELIMINARY; PRT; 979 AA.

AC OSTOK8; 01-FEB-2005 (TREMBLrel. 29, Created)

DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)

DE ENSANGP0000028563 (Fragment).

GN ORFNames=ENSANG0000009133;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;

OX Anopheles; Anopheles.

NX NCBI_Taxid=180454;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.

RT "Anopheles gambiae re-annotation."

RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; ABA01008964; EAL39656.1; Genomic_DNA.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0003725; F:double-stranded RNA binding; IEA.

DR GO; GO:0004519; F:endonuclease activity; IEA.

DR GO; GO:0004525; F:ribonuclease III activity; IEA.

DR GO; GO:0006396; P:RNA processing; IEA.

DR InterPro; IPR001159; DS_RBD.

DR InterPro; IPR000999; RNase_III.

DR Pfam; PF00035; dsm; 1.

DR Pfam; PF00636; Ribonuclease_3; 2.

DR SMART; SM00358; DSRM; 1.

DR SMART; SM00355; RIBOC; 2.

DR PROSITE; PS00137; DS_RBD; 1.

DR PROSITE; PS00137; RNASE_3_1; 1.

DR PROSITE; PS0142; RNASE_3_2; 2.

FT NON TER 1

FT NON TER 1

SQ SEQUENCE 979 AA; 113068 MW; 46F4AE2A0257AD92 CRC64;

Query Match 38.4%; Score 2881.5; DB 2; Length 979;

Best Local Similarity 55.3%; Pred. No. 1.8e-132;

Matches 546; Conservative 169; Mismatches 235; Indels 37; Gaps 12;

```

QY 382 KNYTSIKKEPEETMPDKNEEELIKVWIRCTHSENYSSDPMDQVDSVTVGTSLRL 441
Db 13 RNYCETSEDIARKLEMANDED-----RTLMWRSSPAELIYK-----RVS DKVVESTARKL 62
QY 442 RLVDYKFEELSGROEKAKARPPWPPKTKLDEDESSSESCESEDSTCSSSDSEY 501
Db 63 DALCTLFEELIKRAERITAGTPTVNPBRKXKMKCRKHDKC-----SSSESSEDEEM 117

```



```

QY 502 -PD---VIAEIKKKKAPDRLHDELMWYNDPGQWMDGLCKCSAKARTGIRHSIYPEEA 557
DB 118 EFEDSCSMELBTKIKPRLHVDLWYNDPGEMWMDGLCKCSASRRRTGIRHGICYPEEG 177
QY 558 IKCGRPMNTNAGRLPHYRITVSPPTNFLTDRPTVIEYDDEHYIFEGSMFAHPLNTIPL 617
DB 178 FPCIPSNNAKYLHRIITISPTNFLTITKTDQHEFLBEGSLAHIEGLPT 237
QY 618 CKVIRNIDYTHIFIEEMPEBENCVKGLFSLFLFEDIDELVDMNLKLFEDSPCCP 677
DB 238 CKVIRNIEYTHIIEBQWENFTIRLNFDRFLFELIEVDFTYQ--FSGSGEDSSCP 236
QY 678 RHFMRFRFRFLPDGKEVLSMGIILYLIRCSKALVBEELIAMLQWELNOKTAEBC 737
DB 297 CYHFLRFVFDLPDNGKEVLAMSEVLYLLDNGSLFPLPMLKEMDMMSQWEMODYDVY 356
QY 738 KGMIVNPGTKPSSVHIDOLR-----EQFNPDVITPIIVHIGIRAOISYGDPO 789
DB 357 KGMVNSPGKPKSVARDQLDRNVGDVPEANAIDENLVHVPVLFHGIKPPOLSYAGNPE 416
QY 790 YQKLAKSYVRLHILNANSPKVKOTDKQKLAQREBALQIRKQKTMREVEVLESSOGFMK 849
DB 417 YQKAMEYIKFRILINMNSKSPEDKRLAKENRLLMEMQRMKNITIAVSAKAFHR 476
QY 850 TGRISVDCOHAMMLPVLTTHIRYHQCIMHLDKLGTYEODRCLLQLAMTHPSHHLFGMN 909
DB 477 TGJICMVOHAMLI PVLTHLRFHRSINLVERYIGYFTNRYYTLQALATHPSYKENGCTN 536
QY 910 PDHARNSLSCGIRQPKYGRKVNHMRRKGINTLINISRLGODDPTSRINHERLE 969
DB 537 PDHARNSLTCGIRQPKYGRKIHVMTRKGINTLISIVSRGKEHETDSNITHNERLE 566
QY 970 FLGDVAVEFLTSVHLYLFPSSLEGGIATYRTAIVQNOHLAMLAKKLELDPFMLYAHGPD 1029
DB 597 FLGDVAVEFTISLTHFHPFPDLEGGIATYRTAIVQNOHLAVAKKULHLEFPLYAHGSD 656
QY 1030 LCRESLDRHAMANCFEALIGAVYLESSLEBAKOLFGLFLNDDP-LREVMNLNPLHPLQL 1088
DB 657 LCHELELRHIALNCFEALMGALLLDGIEVADRFAVALFOEDDTLRGIWVNPSPHLOE 716
QY 1089 QEENTRQLETSPLYOKLTFEEBALGVITFHVRLARAFATLVGPNHLLTGHNQMER 1148
DB 717 QEPLGRHNDSEFMKTLTRFEDSIGVQFNHILRLARAFTRDSIGFTNLTLGDSNRLEF 776
QY 1149 LQDSIMQVATEYLFIHPDHEGHLTLRSSLVNNRTOAKVAEBELGMOEYAL--TNDKTK 1207
DB 777 LQDTVQLCSELYLHNFPEHGHSLSSLVNNRTQAVCCDLGAMTOYAYVSNPK-- 834
QY 1208 RPYGLRTKTLADLLESEFIALYTDKOLEVYHTEMVCFPPRLKEFLINQWMDPKSLOQ 1267
DB 835 --ADLTKORADLLEAFGLALYVDKGEYCEMFCHVCLFPLDQFLINQWMDPKSLOQ 892
QY 1268 CCLTLRT-EGKEDDILYKTLQTVGSHARTYVAVYFKGERIGCGKPSIOQAEWGAAM 1326
DB 893 CCLTLRTMGGBDILYVKIETGCTPTNTFVSVAVYFGRKRLACADGSHIOQAEWGAAM 952
QY 1327 DALE--KYNPROMAHQKRFIGKRYROE 1351
DB 953 QALENSKDLFPQDLHQRRVLAQSLKQO 979

```

```

RESULT 12
08CJ74 MOUSE
08CJ74_MOUSE PRELIMINARY; PRT; 541 AA.
AC 08CJ74;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Ribonuclease III.
GN Name=Etch12; Synonyms=Rn3, Rnaasen;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

```

OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Min K.-L., Galanteau A., Parniak M.A.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF440766; AAN63637.1; -, mRNA.
DR HSSP; O67082; 1JFZ.
DR Ensembl; ENSMUSG0000022191; Mus musculus.
DR MGI; MGI:1261425; Etch12.
DR MGI; MGI:1261425; Rnaasen.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003725; F:double-stranded RNA binding; IEA.
DR GO; GO:0004525; F:ribonuclease III activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR InterPro; IPR001159; Ds_RNA_hd.
DR InterPro; IPR000999; Rnaase_III.
DR Pfam; PF00035; dsm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PSS0137; Ds_RBD; 1.
DR PROSITE; PSS0517; RNASE_3_1; 2.
DR PROSITE; PSS0142; RNASE_3_2; 2.
DR SQUENCE 541 AA; 63007 MW; D0AC652A0CE044F1 CRC64;

```

```

Query Match 37.7%; Score 2825; DB 2; Length 541;
Best Local Similarity 98.3%; Pred. No. 5e-130; 6; Indels 0; Gaps 0;
Matches 532; Conservative 3; Mismatches 6;

```

```

QY 834 MREVEVLESSOGFMKTIKRSVCOHAMMLPVLTTHIRYHQCIMHLDKLGTYEODRCLL 893
DB 1 MREVEVLESSOGFMKTIKRSVCOHAMMLPVLTTHIRYHQCIMHLDKLGTYEODRCLL 60
QY 894 QLAMTHPSHHLNFGMNPDHARNSLSCGIRQPKYGRKVNHMRRKGINTLINISRLG 953
DB 61 QLAMTHPSHHLNFGMNPDHARNSLSCGIRQPKYGRKVNHMRRKGINTLINISRLG 120
QY 954 QDDPTSRINHNRLLEFLGDVAVEFLTSVHLYLFPSSLEGGIATYRTAIVQNOHLAMLA 1013
DB 121 QDDPTSRINHNRLLEFLGDVAVEFLTSVHLYLFPSSLEGGIATYRTAIVQNOHLAMLA 180
QY 1014 KKLLEDPFMLYAHGPDLCRESDLRHAMANCFEALIGAVYLESSLEBAKOLFGLFLNDDP 1073
DB 181 KKLLEDPFMLYAHGPDLCRESDLRHAMANCFEALIGAVYLESSLEBAKOLFGLFLNDDP 240
QY 1074 LREVMNLNPLHPLQLQEPNTDRLQLETSPLYOKLTFEEBALGVITFHVRLARAFILRTV 1133
DB 241 LREVMNLNPLHPLQLQEPNTDRLQLETSPLYOKLTFEEBALGVITFHVRLARAFILRTV 300
QY 1134 GFNHLTLGHNRQREFGDSIMQVATEYLFIHPDHEGHLTLRSSLVNNRTOAKVAEB 1193
DB 301 GFNHLTLGHNRQREFGDSIMQVATEYLFIHPDHEGHLTLRSSLVNNRTOAKVAEB 360
QY 1194 LGMQEYAITNDKTKRPVGLRTKTLADLLESEFIALYTDKOLEVYHTEMVCFPPRLKEFI 1253
DB 361 LGMQEYAITNDKTKRPVGLRTKTLADLLESEFIALYTDKOLEVYHTEMVCFPPRLKEFI 420
QY 1254 LNDQWMDPKSLOQCCILTLRTGKEKPDILYKTLQTVGSHARTYVAVYFKGERIGCGK 1313
DB 421 LNDQWMDPKSLOQCCILTLRTGKEKPDILYKTLQTVGSHARTYVAVYFKGERIGCGK 480
QY 1314 GPSIOQAEWGAAMDALAEKYNFPQMAHQKRFIGKRYROELKEMWREBHERBEDTEDIR 1373
DB 481 GPSIOQAEWGAAMDALAEKYNFPQMAHQKRFIGKRYROELKEMWREBHERBEDTEDIR 540
QY 1374 K 1374
DB 541 K 541

```

RESULT 13

Q960Y4 DROME
ID Q960Y4 DROME PRELIMINARY; PRT; 1071 AA.
AC Q960Y4.
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE LD29995p.
GN Name=drosha; ORFNames=CG8730;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Berkeley;
RC Stapleton M., Brocksstein P., Hong L., Agbayan A., Carlson J.,
RA Champé M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
RA Gonzalez M., Guatin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragae V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051775; AAK93199.1; -; mRNA.
DR HSSP; O67082; 1JF2.
DR Ensembl; CG8730; Drosophila melanogaster.
DR Flybase; FBgn026722; CG8730.
DR Flybase; FBgn026722; drosha.
DR GO; GO:0005634; C.nucleus; IEA.
DR GO; GO:0003725; F.double-stranded RNA binding; IEA.
DR GO; GO:0004519; F.endonuclease activity; IEA.
DR GO; GO:0016787; F.hydrolyase activity; IEA.
DR GO; GO:0004525; F.ribonuclease III activity; IEA.
DR GO; GO:0006396; P.RNA processing; IEA.
DR InterPro; IPR001159; D5_RBD.
DR InterPro; IPR000999; RNase_III.
DR Pfam; PF00035; dsm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00355; RIBOC; 2.
DR PROSITE; PS00137; D5_RBD; 1.
DR PROSITE; PS00517; RNase_3_1; 1.
DR PROSITE; PS00142; RNase_3_2; 2.
SQ SEQUENCE 1071 AA; 123158 MW; 25E6C2AADA5078B CRC64;
Query Match 37.5%; Score 2811.5; DB 2; Length 1071;
Best Local Similarity 54.6%; Pred. No. 5.3e-129;
Matches 545; Conservative 169; Mismatches 220; Indels 65; Gaps 15;
QY 412 WTRCTHSENYSSDPMDQVGDSTVGTSTRLRDLYKFESEELSGROEKAKAARPMWEPKPT 471
DB 14 WTRSSPAELYYERTK---SENEVRGRARLQKCTLFDEBLQRAKRVREKLPVYVPPR 69
QY 472 KL-----DEDESSSESCEBDEBOSTCSSSDSEVFDV-----IAETKKAAND 516
DB 70 KARRRVCKHKHSEACSSSSSSDDDEDA-----FKLEQDCMEELSRKVQHQ 119
QY 517 RLHDELWYNDPGQMGDPCKCKAKARTRGIRSHSYGEEBAPRCRPMTNAGRLPHYRI 576
DB 120 RYHADLMHNDAGMGNDGPRCKRCKARRGIRGILPGELGYKLCPPNSNACKLHYRI 179
QY 577 TVSPPTNPLDRPTEVLEDDHEVTFEGFGFMFAAPLPTNPLCKVIRFNI DYTTHFEEMK 636
DB 180 SISPPNPLTKPTTILKDHDEHFLFSGFSLSHVRLSDPLVCVIRFNIETYEYEEKK 239
QY 637 PENFCVKGLELSEFLFRDILLEYDMNLKGPL-----FEDSPCCPFPHMPRVRLPDS 692
DB 240 PENFTIHELDIFPKYLFHELLELVDFNLMPNLPSGNVEES---CPAFHFEPFRVRLPDR 296
QY 693 GKEVLSMHOILLYLAKCSKALVPEBRIANMLQWELMOKVYAECKGMITYNPGTSPSSV 752
DB 297 GKEVLAMVEELARKLQNSAQLVERQQLHNLQISQSEWQYVDFPIKMLATKRGYKPCSI 356
QY 753 RIDQLDREOFNPDV-----ITPPIIVHGIRPAOLSYAGDPQYOKLMSYVXLRLH 803

DB 357 RVDQLDRN--NSDLBECVDRETGISHPAIVHFGICHQPQSLSTAGNEDEYQAKMEYKRYRL 414
QY 804 IANSKRVQTDKOKLAOREEALQKIROKNTMRREYTVELSSQGFMTGIRSDVCOHAML 863
DB 415 MAMMKRPFKDKRKLKEEQRQLEMRTOGRMBRNITVAISSGFRITGIMCVGQAMLI 474
QY 864 PVLTHIRHYOCLMLDKLIGTTPDRCCLQLAMTHPSHHNLFGANPDARNSLSNGIR 923
DB 475 PVLTHIRHFKSLDLLEESIGYFRKRYLLQLALHPSYKEXYGNPDARNSLTNGIR 534
QY 924 QPKYGDRLVHHMHRKKGINTLINIMSRUGODPTPSRINHNERLEFLDAAVEFLTSYH 983
DB 535 QPEYGDRLKHYNTKRGINTLVSTMSRGGKEHETVSNITHERLEFLDAAVEFLTSYH 594
QY 984 LYYLPPSLEEGLATYRTAIVONQHIALAKKLEDPFLYAHGPDLCRESDLRHAMANC 1043
DB 595 LFFMPELEEGSLATYRAIVONQHIALAKKLEDPFLYAHGPDLCHELELRHAMANC 654
QY 1044 FEALIGAVYLEGSLAEAKQLEGRLLF--NDPDLREAVNLVPLHPLQLEPNTROLITSP 1102
DB 655 FEALMGALLLDGIRKADVEFTDALFRODEKLSIMKMLPEHPLQOEFLGDRSCISYR 714
QY 1103 VLQKITERBEALGVFTYVRLARAFPLRTVGFENLTLGHNORMEFLGDSIMQVATEYL 1162
DB 715 VLKELTKREDSIGIFKHIRLARAFTRDSIGFTHLTLGSSNQLRLELGSVTLQICSYL 774
QY 1163 FIFHPDHEGHHTLRSSLVNNRTOAKVAEELGMOBYAI--TYDKTKRPVGLTKTLADLL 1221
DB 775 YHFFEHHEGHLSSLRSSLVNNRTOAVYCDLGMRKXVANYAPK----ADLTKRADLL 830
QY 1222 ESFIALYTDKULEYVHTFMVYCFPRRLKEPLINQDMDNPKSQLOQCCTLTAT--EGKEPD 1280
DB 831 EAFGLALYVDKGLLYCEQCHVCLFPRQLQFLINQDMDNPKSKLQCCCTLTATMDGEBD 890
QY 1281 IPLYKTLQVPSHARTTVANYFEGERIGCGCKGSIQQAEGAMDALEKN--PROVA 1338
DB 891 IPLYKTVASGPNTRVKVAAYFRSKRLATSSGSSIQQAENNAKQALENSRDLPPQD 950
QY 1339 HOKRFYGRKYROEL---KEMRWEREHOEREPDETDIKK 1374
DB 951 HOKRVIANISIKQGTGNELDNDSRQHQ-----EKIKR 983
RESULT 14
Q7PNE9 ANOGA
ID Q7PNE9 ANOGA PRELIMINARY; PRT; 836 AA.
AC Q7PNE9;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000011622 (Fragment).
GN ORFNames=ENSANGG0000009133;
OS Anopheles gambiae str. PE8T.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicidae;
OC Anophelinae; Anopheles.
NCBI_Taxid=180454;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PE8T;
RC The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation";
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PE8T;
RC The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008964; BAA12298.3; -; Genomic_DNA.

DR GO:0005622; C:intracellular; IEA.
DR GO:0003725; F:double-stranded RNA binding; IEA.
DR GO:0004519; F:endonuclease activity; IEA.
DR GO:0004525; F:ribonuclease III activity; IEA.
DR GO:0006396; P:RNA processing; IEA.
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR00999; RNase_III.
DR Pfam: PF00035; dsm1_1.
DR Pfam: PF00636; Ribonuclease_3; 2.
DR SMART: SM00535; RIBOC_2.
DR PROSITE: PS0137; DS_RBD; 1.
DR PROSITE: PS00517; RNase_3_1; 1.
DR PROSITE: PS0142; RNase_3_2; 2.
FT NON TER 1 1
FT NON TER 836 836
SQ SEQUENCE 836 AA; 96229 MW; 8D3E0BD04953877D CRC64;

Query Match 35.8%; Score 2683.5; DB 2; Length 836;
Best Local Similarity 59.2%; Pred. No. 76-123;
Matches 499; Conservative 143; Mismatches 182; Indels 19; Gaps 8;

QY 514 HPRRLDELRYNDPCQNDGFLCKGAKARTGIRHSITPGEALIRCPMTNAGRLTH 573
DB 1 HPRRLHVDLMYNDPGEMNDGFLCRCSARSRTGIRHGKYPGEGFPKCI PNSNNADKLTH 60
QY 574 YRITVSPPTNPLDRPTVIEYDDEHYIPEGFSMFAHAPLNIPLCKVIRPNIDYTHFTE 633
DB 61 YRITVSPPTNPLKPTITLKHQDEHFLPEGFSLAHEPILGELPTCYAIRNIEYTLIYIE 120
QY 634 EMMPENFCVKGLELPSLFLRDILIELYDNLKGPLFEDSPCCRFHMFPRFVRLPDG 693
DB 121 EQMPENFTIRLNLFRVYLFRELELDVFTVQ-PSGSGEDSSPCVHFLPRFRDLPDNG 179
QY 694 KEVLSMHQILLYLIRCSKALVPEEELANMLQWEELEWOKYAECKMITYTNPCTKSSVR 753
DB 180 KEVLSMEVLYRLYLDNSGLVPPDMLKEMWQNMEDVYVYKGVVSNPMGKPCSVR 239
QY 754 IQLDDEGQNPVDITPTIYHFGIRPAQLSYAGDPQYKLMKSYVLYRHLNANSPVKKQT 813
DB 240 VQLDLDNV--GDVPEANALDENG-----YAGNPETQAMREYIYFRHLNANSPSFE 290
QY 814 DRQKLAQREBALOKIRQKNTMRRETVLSSOGFMKTGIRSDVCOHANNLPVLTNHIRYH 873
DB 291 DRKLEAKENRLLEMMQGMKNITITAVSAKAFHGTGIMCDVQDAMLPVLTGHLRTH 350
QY 874 QCLMHLDKLIGYTFQDRCLQLAMTHPSNHLNFGNPDHARNSLSCGIRQPKYGRKYH 933
DB 351 RSLNVLERYIGYFTNRVTLQLALTHPSYKENFGTNPDHARNSLTGCIROPEYGRKIH 410
QY 934 HHMRKKGINTLINISRLGODDPTPSRIINHERLEFLGDVAVEFLTSVHLVYLFPSLE 993
DB 411 YNMTTRRGINTLISIRSGKEHETSNTNHERLEFLGDVAVEFLTSVHLFMPDLDE 470
QY 994 GGLATYRTAIVONQHMLAKKLELDPFMLYHAGPLCRESDRHMANCFEALIGAVY 1053
DB 471 GGLATYRTAIVONQHMLAKKLELDPFMLYHAGSULCHELEKRLANCFEALIGALL 530
QY 1054 ESSLERAKQLFGRLLFNDDP-LAEVNLNPLRPLQLQEPNTDROLISTEVLQKTEFEE 1112
DB 531 DGLIEVADVFAVALFOEDDTLRGIWVYPSHPLQOEPLGDRHNLDSFEMLKTLLTFED 590
QY 1113 AIGVIFTHRLARATLRTVGNNHLLTGNNQMEFLGDSIMQVATVYLFIHPRPHNG 1172
DB 591 SLSGVONHRLRLARATLRTDSTIGFTNLTLGNNQMEFLGSDVLIQISEYLYRFRPHNG 650
QY 1173 HLTLLSSLVNNTQAKVAEELGMOEYAI-TNDKTRPVGLRTKTADLLESFIALYTD 1231
DB 651 HLTLLSSLVNNTQAKVAEELGMOEYAI-TNDKTRPVGLRTKTADLLESFIALYTD 1231
QY 1232 KOLEVYTHVNVCFPRLEKEFLINQDNDPKSQLQCCCLTART-ESEKPDIPLYKTLQTV 1290
DB 707 KGLEVCMECHVCLFPRLODFINQDNDPKSKLQCCCLTARTMDGSEPDIPYKYIECT 766

QY 1291 GSHARTYVAVYFKERICGCKPSIQQAEMGAAMDALE--KYNPQNAHQKRTIGRX 1348
DB 767 GPTNTRVAVYVYFKRKLACADGHSIQQAEMGAAMDALEKNSDLPQLDHQGRVIAQSL 826
QY 1349 RQE 1351
DB 827 KRQ 829

RESULT 15
Q59FP9 HUMAN PRT; 432 AA.
ID Q59FP9 HUMAN PRELIMINARY;
AC Q59FP9;
DC 10-MAY-2005 (TEMBLrel. 30, Created)
DT 10-MAY-2005 (TEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TEMBLrel. 30, Last annotation update)
DE Nuclear RNase III Drosha variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.,
RT "Name Title";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209501; BAD9238.1; -; mRNA.
FT NON TER 1 1
SQ SEQUENCE 432 AA; 50659 MW; E5D81B6DD61251A CRC64;

Query Match 30.6%; Score 2298; DB 2; Length 432;
Best Local Similarity 99.8%; Pred. No. 2,1e-104;
Matches 424; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 556 EAIKPCRPMTNNAAGRLFHYRTVSPPTNPLDRPTVIEYDDEHYIPEGFSMFAHAPLNTN 615
DB 7 QAIKPCRPMTNNAAGRLFHYRTVSPPTNPLDRPTVIEYDDEHYIPEGFSMFAHAPLNTN 66
QY 616 PLCKVIRFNIDYTHFIEEMMPENFCVKGLELPSLFLRDILIELYDNLKGPLFEDSPC 675
DB 67 PLCKVIRFNIDYTHFIEEMMPENFCVKGLELPSLFLRDILIELYDNLKGPLFEDSPC 126
QY 676 CPEFHMPFRVRLPDGKXVLSMHOILYLLRCSKALVPEEELANMLQWEELEWOKYAE 735
DB 127 CPEFHMPFRVRLPDGKXVLSMHOILYLLRCSKALVPEEELANMLQWEELEWOKYAE 186
QY 736 ECKGMVTPNGTKPSSVRIDLDREQNPVDITPTIYHFGIRPAQLSYAGDPQYKLMK 795
DB 187 ECKGMVTPNGTKPSSVRIDLDREQNPVDITPTIYHFGIRPAQLSYAGDPQYKLMK 246
QY 796 SYVKLHLNANSPKVQKQDKLAQREBALOKIRQKNTMRRETVLSSOGFMKTGIRSD 855
DB 247 SYVKLHLNANSPKVQKQDKLAQREBALOKIRQKNTMRRETVLSSOGFMKTGIRSD 306
QY 856 VCOHANNLPVLTNHIRYHOCMLNLDKLIGYTFQDRCLQLAMTHPSNHLNFGNPDHAR 915
DB 307 VCOHANNLPVLTNHIRYHOCMLNLDKLIGYTFQDRCLQLAMTHPSNHLNFGNPDHAR 966
QY 916 SLNSGIRPKYGDRVHHMRKKGINTLINMSRLGDDPTPSRIINHERLEFLGDVA 975
DB 367 SLNSGIRPKYGDRVHHMRKKGINTLINMSRLGDDPTPSRIINHERLEFLGDVA 426
QY 976 VEFLT 980
DB 427 VEFLT 431

Search completed: February 3, 2006, 19:41:12
Job time : 297 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protease - protein search, using sw model

Run on: February 3, 2006, 19:41:24 ; Search time 56 Seconds
(without alignments)
2026.507 Million cell updates/sec

Title: US-10-079-185-2

Perfect score: 7500
Sequence: 1 MMOGNTCHRMSPHGRGCRP.....MRWERHOREPDEDIDK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060. seque, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5.COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/6.COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/H.COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/RE.COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	2	US-09-900-425B-2
2	2463	32.8	466	2	US-09-900-425B-37
3	890	11.9	412	2	US-09-900-425B-3
4	496	6.6	115	2	US-09-513-999C-4635
5	316	4.2	553	2	US-09-513-999C-4635
6	312.5	4.2	241	2	US-09-107-532A-5434
7	284.5	3.8	232	2	US-09-583-110-4554
8	284.5	3.8	236	2	US-09-107-433-4394
9	284.5	3.7	998	2	US-09-949-016-7757
10	277.5	3.7	232	1	US-08-869-674-2
11	277.5	3.7	232	1	US-09-213-010-2
12	277.5	3.7	232	2	US-09-213-011-2
13	271.5	3.6	386	2	US-09-461-774-10
14	267	3.6	243	2	US-08-965-087-2
15	264.5	3.5	1404	2	US-08-801-308-1
16	263	3.5	227	2	US-09-489-039A-11549
17	262	3.5	226	2	US-09-900-425B-6
18	262	3.5	971	2	US-09-248-796A-19531
19	260.5	3.5	245	2	US-09-710-279-1658
20	260.5	3.4	246	2	US-09-134-001C-4787
21	254.5	3.4	607	2	US-09-248-796A-19157
22	252	3.4	228	2	US-09-543-681A-7646
23	248.5	3.3	904	2	US-09-976-594-615
24	239.5	3.2	439	2	US-10-300-819B-21
25	239.5	3.2	865	2	US-09-281-766-19
26	239.5	3.2	865	2	US-09-612-858-19
27	239.5	3.2	865	2	US-09-957-995A-19

28	239	3.2	253	2	US-09-252-991A-29495	Sequence 29495, A
29	233.5	3.1	268	2	US-09-540-236-2003	Sequence 2003, Ap
30	232.5	3.1	1646	2	US-09-535-008-67	Sequence 67, Appl
31	232.5	3.1	1647	2	US-09-535-008-2	Sequence 2, Appl
32	232.5	3.1	1647	2	US-09-824-574-4	Sequence 4, Appl
33	232.5	3.1	1647	2	US-09-538-092-1172	Sequence 1172, Ap
34	232.5	3.1	1649	2	US-09-535-008-75	Sequence 75, Appl
35	232.5	3.1	1650	2	US-09-535-008-71	Sequence 71, Appl
36	232.5	3.1	1659	2	US-09-949-016-9752	Sequence 9752, Ap
37	232.5	3.1	1678	2	US-09-535-008-69	Sequence 69, Appl
38	232.5	3.1	1679	2	US-09-535-008-65	Sequence 65, Appl
39	232.5	3.1	1681	2	US-09-535-008-77	Sequence 77, Appl
40	232.5	3.1	1682	2	US-09-535-008-73	Sequence 73, Appl
41	230.5	3.1	254	2	US-09-533-863-18	Sequence 18, Appl
42	230.5	3.1	1909	2	US-09-590-968B-2	Sequence 2, Appl
43	230	3.1	666	2	US-09-050-739-70	Sequence 70, Appl
44	229.5	3.0	434	2	US-09-538-092-1338	Sequence 1338, Ap
45	228.5	3.0	259	2	US-09-438-185A-56	Sequence 56, Appl

ALIGNMENTS

RESULT 1									
US-09-900-425B-2									
; Sequence 2, Application US/09900425B									
; Patent No. 6737512									
; GENERAL INFORMATION:									
; APPLICANT: Wu, Hongliang									
; APPLICANT: Crooke, Stanley T.									
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof									
; FILE REFERENCE: ISIS029/ISPH-0522									
; CURRENT APPLICATION NUMBER: US/09/900,425B									
; CURRENT FILING DATE: 2001-07-06									
; NUMBER OF SEQ ID NOS: 37									
; SOFTWARE: Patentin version 3.1									
; SEQ ID NO 2									
; LENGTH: 1374									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-900-425B-2									
Query Match 100.0%; Score 7500; DB 2; Length 1374;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MMOGNTCHRMSPHGRGCRPGRGCGHGRPSA	FRPONTLRLHPQDPVVOYQYEPSPAPS	60					
DB	1	MMOGNTCHRMSPHGRGCRPGRGCGHGRPSA	FRPONTLRLHPQDPVVOYQYEPSPAPS	60					
QY	61	TTFSNSPAPVFLPRPDPVFPFPPMPPSAQGLP	PCPIRPPFNHOKRHPVPVPCPPM	120					
DB	61	TTFSNSPAPVFLPRPDPVFPFPPMPPSAQGLP	PCPIRPPFNHOKRHPVPVPCPPM	120					
QY	121	PPMPPCPNPVPVGAAPGQGTFFPMMPPSMPH	PPPPVMPVQVNTQYPPGYSHNNPP	180					
DB	121	PPMPPCPNPVPVGAAPGQGTFFPMMPPSMPH	PPPPVMPVQVNTQYPPGYSHNNPP	180					
QY	121	PPMPPCPNPVPVGAAPGQGTFFPMMPPSMPH	PPPPVMPVQVNTQYPPGYSHNNPP	180					
DB	121	PPMPPCPNPVPVGAAPGQGTFFPMMPPSMPH	PPPPVMPVQVNTQYPPGYSHNNPP	180					
QY	181	SPNSFQNNPSSFLPSANNSSPHFRLLPYPL	LKASERSSPRLKHYDHRDRDHSGR	240					
DB	181	SPNSFQNNPSSFLPSANNSSPHFRLLPYPL	LKASERSSPRLKHYDHRDRDHSGR	240					
QY	241	GERHRLDRRERSRDRRQDSRYRSDVDRGTP	SRHSYERSRERERRRHRDRRS	300					
DB	241	GERHRLDRRERSRDRRQDSRYRSDVDRGTP	SRHSYERSRERERRRHRDRRS	300					
QY	301	PSLERSYKKEYSKSGSYGLSVPEPAGCTPEL	PGELIKNTDMSWAPLEIVNHRSPSREK	360					
DB	301	PSLERSYKKEYSKSGSYGLSVPEPAGCTPEL	PGELIKNTDMSWAPLEIVNHRSPSREK	360					
QY	361	KRARMBEEDKRWSDNOSGDKKNTYSIKKEPE	ETMPDKNEEBEELKVMVIRCHTSEN	420					
DB	361	KRARMBEEDKRWSDNOSGDKKNTYSIKKEPE	ETMPDKNEEBEELKVMVIRCHTSEN	420					

```
QY 421 YSSDPMQVGVSTVGTSLRLDLYDKFEEELSGROEKAAARPMWPKTKLDEBESS 480
| | | | |
DB 421 YSSDPMQVGVSTVGTSLRLDLYDKFEEELSGROEKAAARPMWPKTKLDEBESS 480
QY 481 SESECESDSDSTCSSSDSEVPDVIATKKAHPRRLHDELMTNDPCQNDGSLCKCSA 540
| | | | |
DB 481 SESECESDSDSTCSSSDSEVPDVIATKKAHPRRLHDELMTNDPCQNDGSLCKCSA 540
QY 541 KARRGIRRSIYRGEBAIKPCRPMTNNAARLPFYRTVTSPPNPLFDRPTVIEYDDHEXI 600
| | | | |
DB 541 KARRGIRRSIYRGEBAIKPCRPMTNNAARLPFYRTVTSPPNPLFDRPTVIEYDDHEXI 600
QY 601 FEGFSFMAHAPLTNIPLCKVIRFNIDYTHFIEEMAPENPCVKGLELFLFRDLIELY 660
| | | | |
DB 601 FEGFSFMAHAPLTNIPLCKVIRFNIDYTHFIEEMAPENPCVKGLELFLFRDLIELY 660
QY 661 DNMLKGPLFEDSPCCPRFHEMRFRVRLPDGKEVLSMHOILLYLRSCKALVPEEETA 720
| | | | |
DB 661 DNMLKGPLFEDSPCCPRFHEMRFRVRLPDGKEVLSMHOILLYLRSCKALVPEEETA 720
QY 721 NMLQWEELEWQKYAECEKQIVTNPCTKPSVARDQDREQPNPDVITFPIVHFGIRPA 780
| | | | |
DB 721 NMLQWEELEWQKYAECEKQIVTNPCTKPSVARDQDREQPNPDVITFPIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLMKSYVRLHLLANSPKVKQTDKQKLAQREBAOKIRQKQTMREEVY 840
| | | | |
DB 781 QLSYAGDPQYQKLMKSYVRLHLLANSPKVKQTDKQKLAQREBAOKIRQKQTMREEVY 840
QY 841 ELSSQGFMTKTRISDVCOHAMMLPVLTTHIRYHQCMLHLDKLGTFYFORCLQLAMTHP 900
| | | | |
DB 841 ELSSQGFMTKTRISDVCOHAMMLPVLTTHIRYHQCMLHLDKLGTFYFORCLQLAMTHP 900
QY 901 SHHLNGMNPDBHARNSLNCGRKQPKYGRKVMHMRKKGINTLTINISRLGODDPTSS 960
| | | | |
DB 901 SHHLNGMNPDBHARNSLNCGRKQPKYGRKVMHMRKKGINTLTINISRLGODDPTSS 960
QY 961 RINHNRELEFLGDAVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNOHMLAKKLELDP 1020
| | | | |
DB 961 RINHNRELEFLGDAVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNOHMLAKKLELDP 1020
QY 1021 FMLYAGPRLCRBSDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMIN 1080
| | | | |
DB 1021 FMLYAGPRLCRBSDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMIN 1080
QY 1081 YPLHNPQLOBPNTDROLIETSPVLOKLTBEFEAIGVIFTHVRLARAFLLRTVGFNHLTL 1140
| | | | |
DB 1081 YPLHNPQLOBPNTDROLIETSPVLOKLTBEFEAIGVIFTHVRLARAFLLRTVGFNHLTL 1140
QY 1141 GHNQRMFLGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYA 1200
| | | | |
DB 1141 GHNQRMFLGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYA 1200
QY 1201 ITNDKTRPVGLRTKTLADLLESFIALYTDKOLEVATFPMNCFPRPLKEPFLINDQND 1260
| | | | |
DB 1201 ITNDKTRPVGLRTKTLADLLESFIALYTDKOLEVATFPMNCFPRPLKEPFLINDQND 1260
QY 1261 PKSQLOOCCLTLTBEKEPDIPLKYTLQVGPSHARTYTAAYVFKSGRIGCGGSPSIOQA 1320
| | | | |
DB 1261 PKSQLOOCCLTLTBEKEPDIPLKYTLQVGPSHARTYTAAYVFKSGRIGCGGSPSIOQA 1320
QY 1321 EMGAANDALEKYNPQMAHQKRFIGRKYROELKEMWEREHOBERDEDEDIKK 1374
| | | | |
DB 1321 EMGAANDALEKYNPQMAHQKRFIGRKYROELKEMWEREHOBERDEDEDIKK 1374
```

```
RESULT 2
US-09-900-425B-37
; Sequence 37, Application US/09900425B
; Patent No. 6737512
; GENERAL INFORMATION:
```

```
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
```

```
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425B
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-425B-37
```

```
Query Match 32.8%; Score 2463; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 2,4e-185; Gaps 0;
Matches 466; Conservative 0; Mismatches 0; Indels 0;
```

```
QY 909 NPDHARNSLNCGRKQPKYGRKVMHMRKKGINTLTINISRLGODDPTSPRINHERL 968
| | | | |
DB 1 NPDHARNSLNCGRKQPKYGRKVMHMRKKGINTLTINISRLGODDPTSPRINHERL 60
QY 969 EFLGDAVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNOHMLAKKLELDPMLYAGP 1028
| | | | |
DB 61 EFLGDAVEFLTSVHLVYLPFSLBEGGLATYRTAIVQNOHMLAKKLELDPMLYAGP 120
QY 1029 DLCRBSDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMINPLHPLQ 1088
| | | | |
DB 121 DLCRBSDLRHAMANCFEALIGAVYLEGSLEBAKOLFGRLFPNDPLREVMINPLHPLQ 180
QY 1089 QBPNTDROLIETSPVLOKLTBEFEAIGVIFTHVRLARAFLLRTVGFNHLTIGHNORMEF 1148
| | | | |
DB 181 QBPNTDROLIETSPVLOKLTBEFEAIGVIFTHVRLARAFLLRTVGFNHLTIGHNORMEF 240
QY 1149 LGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTKR 1208
| | | | |
DB 241 LGDSIMQVATEYLFIFHPDHHGHLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTKR 300
QY 1209 PVGLRTKTLADLLESFIALYTDKOLEVATFPMNCFPRPLKEPFLINDQNDPKSQLOOC 1268
| | | | |
DB 301 PVGLRTKTLADLLESFIALYTDKOLEVATFPMNCFPRPLKEPFLINDQNDPKSQLOOC 360
QY 1269 CULTRTEGKEPDIPLKYTLQVGPSHARTYTAAYVFKSGRIGCGGSPSIOAEMGAAMDA 1328
| | | | |
DB 361 CULTRTEGKEPDIPLKYTLQVGPSHARTYTAAYVFKSGRIGCGGSPSIOAEMGAAMDA 420
QY 1329 LEKYNPQMAHQKRFIGRKYROELKEMWEREHOBERDEDEDIKK 1374
| | | | |
DB 421 LEKYNPQMAHQKRFIGRKYROELKEMWEREHOBERDEDEDIKK 466
```

```
RESULT 3
US-09-900-425B-3
; Sequence 3, Application US/09900425B
; Patent No. 6737512
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: ISIS029/ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425B
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-900-425B-3
```

```
Query Match 11.9%; Score 890; DB 2; Length 412;
Best Local Similarity 43.1%; Pred. No. 1,4e-61;
Matches 181; Conservative 85; Mismatches 136; Indels 18; Gaps 7;
```

```
QY 944 TLINISRLGODDPTSPRINHNRELEFLGDAVEFLTSVHLVYLPFSLBEGGLATYRTAIV 1003
```

```
Db      2 SLFNKGTSGGP-----ILHNERLEYLGDVAVVELIVSHLYLPMLTTHHFGGATYRTAL 57
Qy      1004 VONOHIAMLAKKLELDPFMYLAHGPDLCRESDLRHMANCFBALIGAVYLEGSLBEAKOL 1063
Db      58 VONRNIAATLAKNCRIEMLOYSHGADILINAEFKHALANFEVMAIYIDGGLACDVI 117
Qy      1064 FGRLLF-NDPDLREVMANPLHPLLOEPNTDROLIETSPVLOKTEFEBAIVIFTHVR 1122
Db      118 FSAVAGHQPVLEKKMDHINEHLEKREDPGORDLSITFTLTSPHAEERLIDGFNNIR 177
Qy      1123 LLAATTLTGVGNHLLTGNOMERLQDSIMOLVATEYVFIHPDHHSGLTLSSSLV 1182
Db      178 LAKAETRRRIIPNNDLTGHNQRLWLGSDVLQDIVSDFLYRRFPYHNEGHMSLRTSLV 237
Qy      1183 NNRTOAKVAEELGMOEYAITNDKTRPV---GLRTTLDLLESFIAALYTDLDLYVHT 1239
Db      238 SNOTQAVCCDGLFTEFVI-----KAPYKTPBLKLDKADLVBAFGLALYVDHGHICRA 292
Qy      1240 FMNVCFPPRLKEFILNQDNDPKSOLQCCCLTLR-TEGKEPDILYKTLQTVGPHARTY 1298
Db      293 FIRIVCPRLKHRIESEKNNDASHLOQKCLAMRDSSSPDMPEYRVLGIEGPTNNRIF 352
Qy      1299 TAAVYFKGRIGCGKPSIQOAEKGA---AMDLEKYNFPOM-AHQKRFIGRYQELKE 1354
Db      353 KIAVYKGRILASAEASNVKAKELRAVELALANLJESMFSKAKAKNNMNRRLLEDOTSD 412
```

RESULT 4

```
US-09-513-999C-4635
; Sequence 4635, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4635
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23..-1
; OTHER INFORMATION: score 5.9
; OTHER INFORMATION: seq TLKFLTLLOKSNM/KR
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 3
; OTHER INFORMATION: Xaa=Ala or Pro
US-09-513-999C-4635
```

```
Query Match      6.6%; Score 496; DB 2; Length 115;
Best Local Similarity 95.7%; Pred. No. 2, 2e-31;
Matches 88; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      511 KKAHPRLHDELWYNDPGQNDPGLCKCSAKARTGIRHSIYGEBAIKRCRPMYNNAGR 570
Db      24 KRANLRLHDELWYNDPGQNDPGLCKCSAKARTGIRHSIYGEBAIKRCRPMYNNAGR 83
Qy      571 LPHYRTVSPPTNFLTDRPTVYIYDDHEYIFE 602
Db      84 LPHYRTVSPPTNFLTDRPTVYIYDDHEYIFE 115
```

```
RESULT 5
US-09-949-016-7961
; Sequence 7961, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7961
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7961
```

```
Query Match      4.2%; Score 316; DB 2; Length 553;
```

```
Best Local Similarity 28.0%; Pred. No. 3, 2e-16;
```

```
Matches 111; Conservative 32; Mismatches 106; Indels 148; Gaps 21;
```

```
Qy      19 PRGRGHGAPSPASFRPONLRLHLPQDPVOYQYPPSPASTTFNSNPANF-----LPP 74
Db      200 PQGGRGRGPPGA-----VPGGDRFPGPAGP-----CGPPPPPPACQTP 239
Qy      75 RPDVFPFPMPSPAGPLPBCPIRP-----FNNHQMRFHPV-----PPC---PPP 119
Db      240 RP---PLGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 236
Qy      120 MPMPMPCPNPPVPVGAAPGQ-----GTFP-----FMMPSPMHPPP---PPV 159
Db      297 GPP-PPVPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 355
Qy      160 MPQOVYQYPPGYSHNPPPPSPFNQNDPSPFLPANSSSP---HFRHLPY----- 210
Db      356 AP-HVN---PAF-----FPPPT-----NSGMPYSDSGPPPTDPYGRPPYDRGDYG 398
Qy      211 PLPKASERSRSLKHYDDHRRHDS----- 237
Db      399 PPGREMDTARTPLSEAFEFEMNRRAISSAISRAVSDASAGDYSAIETLVTAISLIK 458
Qy      238 -----HG-RGRHRLDRRGRSPDRRQDSRYSDYDRGTP 275
Db      459 QSRVSAADRCVKYLISLQDCLHGISKSYGSGRRERSRDRHSR--SRKSRHKSRSR 516
Qy      276 SRHSYRSRERERHRHRDNRSPSLRSYKEXK 312
Db      517 DRHDYRERSRERHRDRD--RDRDRDRDRRER 551
```

RESULT 6

```
US-09-107-532A-5434
; Sequence 5434, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
```

```
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
```



```

1 ZIP: 02354
2 COMPUTER READABLE FORM:
3 MEDIUM TYPE: CD/ROM ISO9660
4 COMPUTER: PC
5 OPERATING SYSTEM: <Unknown>
6 SOFTWARE: ASCII
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: US/09/107,532A
9 FILING DATE: 30-Jun-1998
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: 60/085,598
12 FILING DATE: 14 May 1998
13 APPLICATION NUMBER: 60/051571
14 FILING DATE: July 2, 1997
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Atinello, Pamela Deneke
17 REGISTRATION NUMBER: 40,489
18 REFERENCE/DOCKET NUMBER: GTC-012
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (781)893-5007
21 TELEFAX: (781)893-8277
22 INFORMATION FOR SEQ ID NO: 5434:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 241 amino acids
25 TYPE: amino acid
26 TOPOLOGY: linear
27 MOLECULE TYPE: protein
28 HYPOTHEICAL: YES
29 ORIGINAL SOURCE:
30 ORGANISM: Enterococcus faecium
31 FEATURE:
32 NAME/KEY: misc feature
33 LOCATION: (B) LOCATION 1...241
34 SEQUENCE DESCRIPTION: SEQ ID NO: 5434:
35
36 US-09-107-532A-5434
37
38 Query Match 4.2% Score 312.5; DB 2; Length 241;
39 Best Local Similarity 33.5%; Pred. No. 1,86-16;
40 Matches 77; Conservative 50; Mismatches 90; Indels 13; Gaps 7;
41
42 Oy 1105 OKLTFEEFAIGVIFTHVRLARAFILRTVGFNH--LTLSHGNRMFPLGDSIMQVATEYL 1162
43 Db 17 QLTRELKKNYNIIVFDLNLLEQAFTHSSVYVNEHRLQLSDNERLFLFDGVADELWVSQYL 76
44 Oy 1163 FIFHPDHHEGHLTLIRSSLVNNRTOAKVAEEIEMOEYATINDKTRPVGRLTK--TLADL 1220
45 Db 77 FRLVPHLPEEGKLTKTTRAAIVREDSLSKFAKECHFDQYVLLG-KGEHNSGGRTRPALLCDL 135
46 Oy 1221 LESPIALAYTKDLEVYHFMNVCFEPRLKEITLQNDNMDPKSQLQQCCCLTRTSGKEPD 1280
47 Db 136 EFAFLGALVLDGQFYTTHAFLEKVIFFPKVKAFAFHEW-DHRTKLOE--VLQKSG---D 188
48 Oy 1281 IPL-YKTLQVGPSPHARTYTVAVYFKGERIGCKGKPSIQDAEMGAMDAL 1329
49 Db 189 VSIETRLINEGPAHERVFWIEVYVDOLIGTGCKSKKLAAQAALENL 238
50
51 RESULT 7
52 US-09-583-110-4554
53 Sequence 4554, Application US/09583110
54 Patent No. 6699703
55 GENERAL INFORMATION:
56 APPLICANT: Lynn Doucette-Stamm et al.
57 TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
58 TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
59 FILE REFERENCE: PATH00-07A
60 CURRENT APPLICATION NUMBER: US/09/583,110
61 CURRENT FILING DATE: 2000-05-26
62 PRIOR APPLICATION NUMBER: US 09/107,433
63 PRIOR FILING DATE: 1998-06-30
64 PRIOR APPLICATION NUMBER: US 60/085,131
65 PRIOR FILING DATE: 1998-05-12
66 PRIOR APPLICATION NUMBER: US 60/051,553

```

/ PRIOR FILING DATE: 1997-07-02
/ NUMBER OF SEQ ID NOS: 5322
/ SEQ ID NO 4554
/ LENGTH: 232
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
/ US-09-583-110-4554

Query Match 3.8%; Score 284.5; DB 2; Length 232;
Best Local Similarity 32.9%; Pred. No. 2.8e-14;
Matches 76; Conservative 48; Mismatches 90; Indels 17; Gaps 8;

Dy 1108 TEFEEAIGVITFTHVALLARAFTLRTVGFNH--LTGHNQRMFEFGDSINQLVAITYEYFIH 1165
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 6 TVLNKHFAPFBETDKKLLETAFTHTSYANEHRLTKTSHNERLEFLGDAVIQLLSLEYLYKK 65

Dy 1166 FPDHHEGLTLIRSSLVNNRTOAKVAEELGMGEYAITNKTGRPVGLRTK--TLADLLS 1223
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 66 YPKREGLSKLRAMIVAREESLAGFARDCCQDFQF-IKLKGEEKSGGRNRDITLGDAFEA 124

Dy 1224 FIALLTYDQDLLEVYTFPMNVCFPRLK--EPIIDPNMDPKSQLQOCCULTRTGKEPDI 1281
 | : | | | | : | : | : | : | : | : | : | : | : | : | : | : | :
 | : | | | | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 125 FLGALLLDKDAAKVAFELYQMIPKVEAGEFEWITDY--KTHLOE--LLQNG--DV 175

Dy 1282 PL-YKTIQTGVPSHARTYTVAVYFKGERIGCGKSPSIQAEMGANPADEAK 1331
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 176 AIRQVISETGPAMDKVFDEVLVEGKSIGQGGSKSLAEQRAAKNAVER 226

RESULT 8
US-09-107-433-4394
Sequence 4394, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4394:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEtical: YES

```

? ORIGINAL SOURCE:
? ORGANISM: Streptococcus pneumoniae
?
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (B) LOCATION 1..236
? SEQUENCE DESCRIPTION: SEQ ID NO: 4394:
US-09-107-433-4394

```

Query Match	3.8%	Score 284.5	DB 2	Length 236
Similarity	32.9%	Pred. No. 2.8e-14		
Best Local				
Matches 76	Conservative 48	Mismatches 90	Indels 17	Gaps 8

[illegible]

```

RESULT 9
US-09-949-016-7757
: Sequence 7757, Application US/09949016
: Patent No. 661239
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7757
: LENGTH: 998
: TYPE: PRT
: ORGANISM: Human
: US-09-949-016-7757

```

Query March	3.8;	Score 284.5;	DB 2;	Length 998;
Best Local Similarity	22.7;	Pred. No. 2.3e-13;		
Matches 151; Conservative	63;	Mismatches 206;	Indels 245;	Gaps 30;

Qy	27	ARPEA--PSFPPQLRLHPCQRPQVGYEERPSA-----PPTTSNSDAPAFLEPR	75
Dh	350	ARRPGGRRPGWHSN-KLGVLVSPQQILRGERSCYRSINRGRIHNSQRQGPB-LPAT	407
Qy	76	PDFFPPPPMWPMSAQGRPLRPPCIRPPPPNHHMRHFPVPCFPMPRPMPCSPNNPPVGA	135
Dh	408	PVFVPVPPVPP-----PLVPPPP-----HTLRPL--PGVPPPPQFSGPFP--PGQ	445
Qy	136	PPGGGTPPEMMPPSPMHPRPVPPVMPQVN-----YQYP-----	170
Dh	446	PPPAG---YSPPPGFP-PAPANLSTPMVSSGVGTAAHSNTLPTQAPRLSREEFFREQR	501
Qy	171	-----GYSHNFPPTSPNS	184
Dh	502	LKEEKKKKSKLDEFTNDFAKELMEYKKIQEKRRRSFSKSKSPYSGSSYSRSLTYYSKRS	561

```

Oy 185 FQNNSSFLBANNSSPHFRLPPYPLPKA PSEERSPEBLKHYDHRHDSHGGERH 2444
Db 562 GSTRSSYSRFSFSRSHSRYSRSPYP-----RKGKSRNY--RSHRSHG--YH 608
Oy 245 RSLDRERGRSPDRRDSRYNSDDV-RGTPSRHRSYERSREER----- 2899
Db 609 RS-----RSRSPYRRYHSRSRSPQAFGQSPKNKNVPGETEREVFNRYREVPYDMK 663
Oy 290 -----EHRHRDRRSFSLRSYKKEYKSGRGYSVVPBPAQCTPELPG 3366
Db 664 AYYGRSVDRDPPEKERYE-----WERYKREWEEKYKAYAAQPRPSA----- 7099
Oy 337 IIKNTDSMAP-----PLEIVNHRSPREKKRARWEEKRMDNDSG-----KD 3811
Db 710 ---NRNFPERFLPNTIRN--SPTRRRRDRDYVGQGHRSRNRGINSYPEKLSARDHQ 7646
Oy 382 KNTYSIKKEPBEYWD-----KOEBEELIKVWIRCTH-----SENY 4211
Db 765 KDMTKSEKESBNAPDQGNKHKRKRGKEEBEGLNELLTETSRKSHPTVEENK 8242
Oy 422 YSS-----DPMQVDSYVGTSLRDLJDYDFEEL-----GSRQ 4566
Db 825 TBSLFLPILBRDAPVPRBPMV--AESITTFYSVSEKDGREGRDKAKAGDKTKRNDSSAV 8822
Oy 457 EKAKAAPWPEPPKTKLDELDSSSESC-----ESDEDSTCSSSDSEVDFVIAEIK 5099
Db 883 SKKENIVKPAKGQEKVDGERERSRSPPIKAKAEPTKDTNKTSSSSSQKOEKINGTP 9422
Oy 510 RKTAK 514
Db 943 R-KAH 946

```

RESULT 10
 US-08-869-674-2
 Sequence 2, Application US/08869674
 Patent No. 5865365
 GENERAL INFORMATION:
 APPLICANT: Lometto, Michael A.
 APPLICANT: Rosenberg, Martin
 TITLE OF INVENTION: NOVEL COMPOUNDS
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dechart Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NT
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/869,674
 FILING DATE:
 CLASSIFICATION: 436
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: GM10013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 232 amino acids

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,308
FILING DATE: 18-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 372,643P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-308-1

Query Match 3.5%; Score 264.5; DB 2; Length 1404;
Best Local Similarity 19.9%; Pred. No. 1,4e-11;
Matches 214; Conservative 125; Mismatches 327; Indels 411; Gaps 50;

QY 25 HGARPSA--PSFRPQNLRLHPQOQPVQYQYEPSPA-----DSTFTSSNPAPNPLP 73
DB 314 NNAAPCGGRPGWHSN-KLGYLVSPQQLRGERSCYRSINRGHHSERSQRTQSPS-LP 371
QY 74 PRPDVPPPPPPPPSAQGLPFCPIRPPPNHQMHPFVPPCFPPMPPPMCPNNPPVP 133
DB 372 ATPCFVPVPP-----PPLYPPPPHLPPLP-----P 396
QY 134 GAPPGCGT--PPMPMPMPHPPPPPVMPQVNYQ--YPPGY--SHNFPSPS----- 181
DB 397 GVPPOFSGPQFPSSQPTAGYSVPPPGPPAPANISTACPSGVPPTAHNSMTPTQAPLL 456
QY 182 ----FNSFQNNP-----SSFPLPSANNS-----SPHFRH----LPPY 210
DB 457 SREEFRQNDKGRSKFPYSGSYSSSYTDSQGLAQHIALTSPSAATLIDLHDH 516
QY 211 PLPKASERSRSPRLGHYDHRHDHSHGGERHNRSLDRERGRSPDRRQDSRYSDYD 270
DB 517 PHPPEAEAR--AMIVHMPDLMDIAH-----KRSRSPYRRYRSRSRSPPE 561
QY 271 -RGRTSPRHSYERSRERER-----ERHRHNRNRSRS 302
DB 562 FRGQSTPKANVPREEKEREYFNRIYREVPPIYDIKAYGRSVDRDPFEKERIYE----- 615
QY 303 LERSYKKEYKRSGRSYGLSVPEPACGTBELPGEIINKTDSWAP---PLEIVNHRSPSR 358
DB 616 WERKYREWYKYGAVGAPRPSA-----NREDFSPRLPLFINRN--SPFT 662
QY 359 EKKRAWEEKORWSDN-----QSSGCKQNTYTSIKEKEPEETMPD----- 398
DB 663 RGRREDYAGQSHRNRLGNYPEKLSRDSHNAKDNPKSKESSENVPDGGKGNHKKR 722
QY 399 ----KNE---EBEBELKPVMI-----RC-----THSENY--SSDPMDQVGDSTVVG 437
DB 723 RKRNEKEGEBESFLNPBELLETSRKCRGSSGIDETKTIDLFLVPSRDDATPVADSPMDA 782
QY 438 TS-RLDLYDKFEEL-----GSRQEKAKAAPWPWPPTKLDLEDELESS 481
DB 783 EGIITPFSVSDKDKREKDKPVKSDKTKRKSDGSAATAKDNVLKPSKGPQEKVDGDEKSP 842
QY 482 ESE---CESDEDT-----CSSSDSEVFVYIAEIKKKAHPRLHDELMTYNDPGQM 530
DB 843 RSEPPPLKKAKEEATKIDSVPSSSSQDEKVTGT-----PRKAHSKSAKQTRRSQPR-- 895
QY 531 NDGPLCKSAKARTGIRHSIYGEBAIKRCRPMTNAGRLPHYRTVTSPTNFLTDRPT 590
DB 896 -----TRSRKTVPKTS---SQKSGPVRTTRPPSLAKINY--LIAEKNEREKRK 941

QY 591 VIEYDDHEVIFEGFSMPAHAPLTNIPLCVIRFNIDYTHFLHEMMPEMFCYKGLFSL 650
DB 942 SYDKD-----FESSSM-----KISK-----VEGTETVXP 965
QY 651 FLFR-----DILEYDMNLKGPLFEDSPCCFRHFMPFRVRLPDGKEV----- 696
DB 966 SPKKMGEDVEKLETPPEKDKIASSTPA-----KKIKLNREIGKIGNAENASTTK 1017
QY 697 -----LSMHQILLYLLR-----CSKALV----- 714
DB 1018 EPSEKLESTSKIKQEKVYGAKRKVAGSESSSTLVDTYSTSGSPVRKSEKTDYK 1077
QY 715 -----PEREIANMIQ-----WELDEMOKYAECK-----GMTV 742
DB 1078 RTVIKIMEEYNDNTAPAEVDVITIMIQVPSKWKDKDFESEEDVKTTPRIQSVGKPSII 1137
QY 743 TNPGTKPSSVRIDQDREQFNPDVITFPIIVHFGIRPAQLSYAGDPQYQKWK--SYVRL 800
DB 1138 KAVTTKPSKT-AKYTEKSEQPS-----KLOKLPKEASHHELM 1173
QY 801 RHLL-----ANSPKVKQTDKQKLAQREBALQKIROKNTMRREVTVE-LSSQGFWKT 850
DB 1174 QHELRSSKGSASSEKGRADREHSGSEKONPDKRKSQAQPDKESTVDRLSEOGHFKT 1230

Search completed: February 3, 2006, 19:43:08
Job time : 59 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2006, 19:42:19 ; Search time 200 Seconds
(without alignments)
2870.487 Million cell updates/sec

Title: US-10-079-185-2

Sequence: 1 MMOGNTCHRMSPHFGRCGCRP.....MRMRERHOREPDEDIDK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published Applications_AA_Main:*
2: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7500	100.0	1374	3	US-09-900-425A-2
2	7500	100.0	1374	4	US-10-079-185-2
3	7500	100.0	1374	4	US-10-774-974-2
4	7500	100.0	1374	4	US-10-805-919-2
5	7500	100.0	1374	6	US-11-001-993-2
6	7486	99.8	1374	5	US-10-723-860-281
7	7486	99.8	1374	5	US-10-756-149-4809
8	2963	39.5	1327	6	US-11-097-143-2409
9	2847	38.0	541	4	US-10-205-331-116
10	2463	32.8	466	4	US-10-774-974-37
11	1961	26.1	378	4	US-10-103-313-471
12	1320	17.6	263	4	US-10-103-313-307
13	890	11.9	412	3	US-09-900-425A-3
14	890	11.9	412	4	US-10-079-185-3
15	890	11.9	412	4	US-10-774-974-3
16	890	11.9	412	4	US-10-805-919-3
17	890	11.9	412	6	US-11-001-993-3
18	340	4.5	1151	6	US-11-097-143-11586
19	311	4.1	560	5	US-10-450-763-51399
20	309.5	4.1	1557	4	US-10-437-963-137427
21	303.5	4.0	406	5	US-10-450-763-57609
22	298.5	4.0	237	4	US-10-437-963-130873
23	297.5	4.0	800	4	US-10-029-386-32198
24	292	3.9	466	4	US-10-437-963-167964
25	287.5	3.8	229	4	US-10-424-599-207316
26	285.5	3.8	258	4	US-10-437-963-159481
27	284.5	3.8	236	5	US-10-617-320-4394

28	284.5	3.8	1616	5	US-10-934-998-88	Sequence 88, Appl
29	282.5	3.8	445	4	US-10-437-963-161789	Sequence 161789,
30	282	3.8	329	4	US-10-437-963-197333	Sequence 197333,
31	281	3.7	598	4	US-10-282-122A-64954	Sequence 64954, A
32	280	3.7	266	4	US-10-425-115-189501	Sequence 189501,
33	279.5	3.7	232	5	US-10-472-928-2508	Sequence 2508, Ap
34	278	3.7	180	4	US-10-425-115-229737	Sequence 229737,
35	277.5	3.7	285	4	US-10-437-963-189482	Sequence 189482,
36	277	3.7	1537	6	US-11-097-143-27759	Sequence 27759, A
37	276.5	3.7	187	4	US-10-425-115-352789	Sequence 352789,
38	276	3.7	273	4	US-10-425-115-330708	Sequence 330708,
39	276	3.7	948	4	US-10-308-448-15	Sequence 15, Appl
40	276	3.7	948	5	US-10-934-998-295	Sequence 295, App
41	275.5	3.7	521	4	US-10-437-963-179715	Sequence 179715,
42	275.5	3.7	594	6	US-11-097-143-10878	Sequence 10878, A
43	274.5	3.7	399	4	US-10-437-963-124007	Sequence 124007,
44	274	3.7	280	4	US-10-425-115-231956	Sequence 231956,
45	273.5	3.6	431	4	US-10-437-963-204963	Sequence 204963,

ALIGNMENTS

```
RESULT 1
US-09-900-425A-2
; Sequence 2, Application US/0900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-425A-2

Query Match      100.0%; Score 7500; DB 3; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MMOGNTCHRMSPHFGRCGCRGCGARPSAPSPFRQNTLRLHPQOPVVOYQYEPSAPS 60
        |||
DB      1 MMOGNTCHRMSPHFGRCGCRGCGARPSAPSPFRQNTLRLHPQOPVVOYQYEPSAPS 60
        |||

QY      61 TTFSNSPAPNPLPPRPDPVPPPPMPSPAGPPLPPCPILPPFPNNHQRHPPPPCPPPM 120
        |||
DB      61 TTFSNSPAPNPLPPRPDPVPPPPMPSPAGPPLPPCPILPPFPNNHQRHPPPPCPPPM 120
        |||

QY      121 PRPMPCPNPVPVGAAPGCGTFPFMPPPSPMPHPPPVPMQOVNVOYQYPGYSHNFPPP 180
        |||
DB      121 PRPMPCPNPVPVGAAPGCGTFPFMPPPSPMPHPPPVPMQOVNVOYQYPGYSHNFPPP 180
        |||

QY      181 SFNSFQNNSSSFLPSANNSSSPHFRLLPYPLPKASERRSPERLKGYYDDHRHDSHGR 240
        |||
DB      181 SFNSFQNNSSSFLPSANNSSSPHFRLLPYPLPKASERRSPERLKGYYDDHRHDSHGR 240
        |||

QY      241 GGHRSGLDERRGRSPDRRQDSRYRSDYDRGTPRRHSYRSRRERBRHRHRNRRS 300
        |||
DB      241 GGHRSGLDERRGRSPDRRQDSRYRSDYDRGTPRRHSYRSRRERBRHRHRNRRS 300
        |||

QY      301 PSLERSYKKEYSKRGSGYGLSVVPEPAGCTPELPGELIKNTUSWAPLEIVNHRSPSRER 360
        |||
```

```
|||||
Db 301 PSLERSYKKEYSRSGSYGLSVPEPAGCTPELPGIINKNTDSWAPLEIVNHSRSRK 360
Qy 361 KARAMEEKDRMSDNOSSGDKKNYTSIKEKEPETPMDKNEEBEELKPVWIRCHSEN 420
Db 361 KARAMEEKDRMSDNOSSGDKKNYTSIKEKEPETPMDKNEEBEELKPVWIRCHSEN 420
Qy 421 YSSDPMQVGSSTVVGTSRLDLVDKFEELSGROEKAARPMPEPKTKLDEDESS 480
Db 421 YSSDPMQVGSSTVVGTSRLDLVDKFEELSGROEKAARPMPEPKTKLDEDESS 480
Qy 481 SSECESDSDSTCSSSDSEVPVIAEIKRKAHPRLHDELMYNDPGQNDGRLCKCSA 540
Db 481 SSECESDSDSTCSSSDSEVPVIAEIKRKAHPRLHDELMYNDPGQNDGRLCKCSA 540
Qy 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAARLPHYRTVSPPTNPLTDRPTVIEYDDHEI 600
Db 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAARLPHYRTVSPPTNPLTDRPTVIEYDDHEI 600
Qy 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTIHFIEEMMPENFCVKGLEFSLFPRDILEY 660
Db 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTIHFIEEMMPENFCVKGLEFSLFPRDILEY 660
Qy 661 DNMLKGPLFEDSPCCPRPHFMRFPVRLPDGKEVLSMHOIILYLRSKALVPEEETA 720
Db 661 DNMLKGPLFEDSPCCPRPHFMRFPVRLPDGKEVLSMHOIILYLRSKALVPEEETA 720
Qy 721 NMLQWELMOKYAECECKMIVTNPGTKPSSVARDLDREOFNPDUVTPIYHFGIRPA 780
Db 721 NMLQWELMOKYAECECKMIVTNPGTKPSSVARDLDREOFNPDUVTPIYHFGIRPA 780
Qy 781 QLSYAGDPQYOKLWKSYYVLRHLANSPKVKQTDKQKLAOREBALOKIRQKMTMRRETV 840
Db 781 QLSYAGDPQYOKLWKSYYVLRHLANSPKVKQTDKQKLAOREBALOKIRQKMTMRRETV 840
Qy 841 ELSSSGCFMKTGIRSDVCOHAMMLPVLTNHIRYQCIMHLDKIGYFQDRCLQLAMTPR 900
Db 841 ELSSSGCFMKTGIRSDVCOHAMMLPVLTNHIRYQCIMHLDKIGYFQDRCLQLAMTPR 900
Qy 901 SHHLNGKMPDHARNSLSNCGIRQPKYGDARKVHNMKKGINTLNINSRLGODDPTS 960
Db 901 SHHLNGKMPDHARNSLSNCGIRQPKYGDARKVHNMKKGINTLNINSRLGODDPTS 960
Qy 961 RINHERLEFLGDAVEFLTSVHLUYLFPSLBEGSLATYRTAIVQNHMLAMKLLEDP 1020
Db 961 RINHERLEFLGDAVEFLTSVHLUYLFPSLBEGSLATYRTAIVQNHMLAMKLLEDP 1020
Qy 1021 FMLYAHGPRLCRESDLRHAMANGFELLIGAUVLEGSLEBAKOLFGRLFNDRDLREVMEN 1080
Db 1021 FMLYAHGPRLCRESDLRHAMANGFELLIGAUVLEGSLEBAKOLFGRLFNDRDLREVMEN 1080
Qy 1081 YRLHPRLOLOBPNDROLIETSPVLOKLTPEFEAIGVIFTHVRLAARFLLTGAFNHLTL 1140
Db 1081 YRLHPRLOLOBPNDROLIETSPVLOKLTPEFEAIGVIFTHVRLAARFLLTGAFNHLTL 1140
Qy 1141 GHNORMEFLGDSIMOLVATEYLFIHPPDHNBHGLTLRSLSLVNRTQAQVAEELGMOEYA 1200
Db 1141 GHNORMEFLGDSIMOLVATEYLFIHPPDHNBHGLTLRSLSLVNRTQAQVAEELGMOEYA 1200
Qy 1201 ITNDKTRPVGLRTKTLADLLESFIALALYTDKOLEVYHFMNVCFFPRKEFLINDAMD 1260
Db 1201 ITNDKTRPVGLRTKTLADLLESFIALALYTDKOLEVYHFMNVCFFPRKEFLINDAMD 1260
Qy 1261 PMSOLOCCCLLTATEGKEBDPIPLYKTLQVGSHPARTYAAVFKBERIGCGGSPSIOQA 1320
Db 1261 PMSOLOCCCLLTATEGKEBDPIPLYKTLQVGSHPARTYAAVFKBERIGCGGSPSIOQA 1320
Qy 1321 EMGAAMDALBKYNPOMAHOKRFIGRKYROELKEMMEREHOREDEDEDIKK 1374
Db 1321 EMGAAMDALBKYNPOMAHOKRFIGRKYROELKEMMEREHOREDEDEDIKK 1374
```

```
US-10-079-185-2
; Sequence 2, Application US/10079185
; Publication No. US2003004941A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: IS155030
; CURRENT APPLICATION NUMBER: US/10/079,185
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-079-185-2

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
1 MNOGNTCHMSFHPRGCGRGHGHGAPSAFSPONRLHPOOPVQOYEPPSAS 60
|||||
Db 1 MNOGNTCHMSFHPRGCGRGHGHGAPSAFSPONRLHPOOPVQOYEPPSAS 60
Qy 61 TTFSNSPAPNLFPRPDPVFPVPPMPPSAOGFLPCCIPRPPPNHQRAPFVPPCFPPM 120
Db 61 TTFSNSPAPNLFPRPDPVFPVPPMPPSAOGFLPCCIPRPPPNHQRAPFVPPCFPPM 120
Qy 121 PPPMPCPNPVPYGAAPPGOGTFPPMPPPSMHPPPPPMPOQVYQYPPGISHNFP 180
Db 121 PPPMPCPNPVPYGAAPPGOGTFPPMPPPSMHPPPPPMPOQVYQYPPGISHNFP 180
Qy 181 SFNSFONNPSFLPANSSSPHFRHLPLPYLPKAPSEERSPERLKHVDHHRDHSOR 240
Db 181 SFNSFONNPSFLPANSSSPHFRHLPLPYLPKAPSEERSPERLKHVDHHRDHSOR 240
Qy 241 GERHSLDRERGRSPDRRQDSRYRSDYDRGRTSPRHSYERSERERERHRHNRSS 300
Db 241 GERHSLDRERGRSPDRRQDSRYRSDYDRGRTSPRHSYERSERERERHRHNRSS 300
Qy 301 PSLERSYKKEYSRSGSYGLSVPEPAGCTPELPGIINKNTDSWAPLEIVNHSRSRK 360
Db 301 PSLERSYKKEYSRSGSYGLSVPEPAGCTPELPGIINKNTDSWAPLEIVNHSRSRK 360
Qy 361 KARAMEEKDRMSDNOSSGDKKNYTSIKEKEPETPMDKNEEBEELKPVWIRCHSEN 420
Db 361 KARAMEEKDRMSDNOSSGDKKNYTSIKEKEPETPMDKNEEBEELKPVWIRCHSEN 420
Qy 421 YSSDPMQVGSSTVVGTSRLDLVDKFEELSGROEKAARPMPEPKTKLDEDESS 480
Db 421 YSSDPMQVGSSTVVGTSRLDLVDKFEELSGROEKAARPMPEPKTKLDEDESS 480
Qy 481 SSECESDSDSTCSSSDSEVPVIAEIKRKAHPRLHDELMYNDPGQNDGRLCKCSA 540
Db 481 SSECESDSDSTCSSSDSEVPVIAEIKRKAHPRLHDELMYNDPGQNDGRLCKCSA 540
Qy 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAARLPHYRTVSPPTNPLTDRPTVIEYDDHEI 600
Db 541 KARRTGIRHSIYPGEBAIKPCRPMTNNAARLPHYRTVSPPTNPLTDRPTVIEYDDHEI 600
Qy 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTIHFIEEMMPENFCVKGLEFSLFPRDILEY 660
Db 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTIHFIEEMMPENFCVKGLEFSLFPRDILEY 660
```


QY 661 DNMLKGPFLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRSKALVPEBEIA 720
DB 661 DNMLKGPFLFEDSPCCPRFHFMRFVRFLPDGKEVLSMHQILLYLRSKALVPEBEIA 720
QY 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIIDQIDREQFNPDIYTFPIIVHFGIRPA 780
DB 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIIDQIDREQFNPDIYTFPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLKMSYVYKLRIHLANS PKVKQTDKOKLAQREBALOKIROKNTMRREVTY 840
DB 781 QLSYAGDPQYQKLKMSYVYKLRIHLANS PKVKQTDKOKLAQREBALOKIROKNTMRREVTY 840
QY 841 ELSSQGFMTGTGIRSDVCOHAMMLPVLTHTIRYHQCLMHLDKLIGYTFQDRCLQLAMTHP 900
DB 841 ELSSQGFMTGTGIRSDVCOHAMMLPVLTHTIRYHQCLMHLDKLIGYTFQDRCLQLAMTHP 900
QY 901 SHHLNFGKMPDHAARNLSNCGIRQPKYGDRAKHHMMRKKGINTLTINISRLQODDPTPS 960
DB 901 SHHLNFGKMPDHAARNLSNCGIRQPKYGDRAKHHMMRKKGINTLTINISRLQODDPTPS 960
QY 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGGLATYRTAIVQNOHMLAMAKKLELD 1020
DB 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGGLATYRTAIVQNOHMLAMAKKLELD 1020
QY 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGA VLEGSLEBAKOLFGRLLFNDPDLREVWLN 1080
DB 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGA VLEGSLEBAKOLFGRLLFNDPDLREVWLN 1080
QY 1081 YPLHPIQLOEPNTDROLIESPVLOKLTFFEEAIGYITFHVRLAAAFITRTYGFNHLTL 1140
DB 1081 YPLHPIQLOEPNTDROLIESPVLOKLTFFEEAIGYITFHVRLAAAFITRTYGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMQVATETYLFIHFPDHHGHLTLRSLSLVNRTQAKAYEELGMOEYA 1200
DB 1141 GHNORMEFLGDSIMQVATETYLFIHFPDHHGHLTLRSLSLVNRTQAKAYEELGMOEYA 1200
QY 1201 ITNDTKRPVGLATKTTLADLESFIALYTDKOLEYVHTFMVCFPPRLKEPILNDQMD 1260
DB 1201 ITNDTKRPVGLATKTTLADLESFIALYTDKOLEYVHTFMVCFPPRLKEPILNDQMD 1260
QY 1261 PXSQLOOCCLTLTTEKEKRPILYKTLQTVGSHATYTYAVYFKGBRIGCGGSPSIOQA 1320
DB 1261 PXSQLOOCCLTLTTEKEKRPILYKTLQTVGSHATYTYAVYFKGBRIGCGGSPSIOQA 1320
QY 1321 BNGAAMDAL EKVPOMAHOKRPIGRKYGOELKEMMEREHOREDEDEDIKK 1374
DB 1321 BNGAAMDAL EKVPOMAHOKRPIGRKYGOELKEMMEREHOREDEDEDIKK 1374

RESULT 3

US-10-774-974-2
/ Sequence 2, Application US/10774974
/ Publication No. US20040126867A1
/ GENERAL INFORMATION:
/ APPLICANT: Wu, Hongjiang
/ APPLICANT: Crooke, Stanley T.
/ TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
/ FILE REFERENCE: ISIS5029/ISPH-0522
/ CURRENT APPLICATION NUMBER: US/10/774, 974
/ PRIOR FILING DATE: 2004-02-09
/ PRIOR APPLICATION NUMBER: US/09/900, 425B
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 2
/ LENGTH: 1374
/ TYPE: PR
/ ORGANISM: Homo sapiens
US-10-774-974-2

Query Match 100.0%; Score 7500; DB 4; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMOGNTCHNRSFHPGCGPRGCGHGAAPSAPSRPQNTLLHPQOPVQOYQYEPSAPS 60
DB 1 MMOGNTCHNRSFHPGCGPRGCGHGAAPSAPSRPQNTLLHPQOPVQOYQYEPSAPS 60
QY 61 TTFSSNAPNFFLPFRDPFVFPFPPMPSAQPLPFCPTIAPPFNHOMRHPFPFPCFPFM 120
DB 61 TTFSSNAPNFFLPFRDPFVFPFPPMPSAQPLPFCPTIAPPFNHOMRHPFPFPCFPFM 120
QY 121 PRPMPCNNPVPVGAAPGQTFPPMMPPPSMHPAPPVPMQOVNTQYPPGYSHNHPPP 180
DB 121 PRPMPCNNPVPVGAAPGQTFPPMMPPPSMHPAPPVPMQOVNTQYPPGYSHNHPPP 180
QY 181 SFNSPQNNSSFLPSANSSSPHFRHLPPYPLPKASERRSPERLKHYDHRHRDHSGR 240
DB 181 SFNSPQNNSSFLPSANSSSPHFRHLPPYPLPKASERRSPERLKHYDHRHRDHSGR 240
QY 241 GERHRLDRRERGRSDRRRODSRYSDYDRGTPSRHSYERSERERHHRDNRRS 300
DB 241 GERHRLDRRERGRSDRRRODSRYSDYDRGTPSRHSYERSERERHHRDNRRS 300
QY 301 PSLESYKKKEXKSGSYGLSVVPEPAGCTPBLPGELIKNTDSWAPLEIVNHRSPSREK 360
DB 301 PSLESYKKKEXKSGSYGLSVVPEPAGCTPBLPGELIKNTDSWAPLEIVNHRSPSREK 360
QY 361 KRARWEEKDRMSDNOSSGCKKNYTSIKEKEPEETVPDKNEEBEELKPVWIRCTHSEN 420
DB 361 KRARWEEKDRMSDNOSSGCKKNYTSIKEKEPEETVPDKNEEBEELKPVWIRCTHSEN 420
QY 421 YSSSDMDQVGDSTVGTGRRLDLYKFEBELGROEKAAPRPMPEPKTKLDEBLESS 480
DB 421 YSSSDMDQVGDSTVGTGRRLDLYKFEBELGROEKAAPRPMPEPKTKLDEBLESS 480
QY 481 SSESCESDSDSTCSSDSEVPVIAIRKKKAPRLDELYNPGQNDGPLCKCSA 540
DB 481 SSESCESDSDSTCSSDSEVPVIAIRKKKAPRLDELYNPGQNDGPLCKCSA 540
QY 541 KARTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTIVSPPTNFLTDRPTVIEYDHEYI 600
DB 541 KARTGIRHSIYGEBAIKPCRPMTNAGRLPHYRTIVSPPTNFLTDRPTVIEYDHEYI 600
QY 601 PEGSMPFAVAPLTNIPLCVYIRFNIDYTHFIEEMMPENVCVGLFSLFRDILELY 660
DB 601 PEGSMPFAVAPLTNIPLCVYIRFNIDYTHFIEEMMPENVCVGLFSLFRDILELY 660
QY 661 DNMLKGPFLFEDSPCCPRFHFMRFVRFLPDGKETLSMHQILLYLRSKALVPEBEIA 720
DB 661 DNMLKGPFLFEDSPCCPRFHFMRFVRFLPDGKETLSMHQILLYLRSKALVPEBEIA 720
QY 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIIDQIDREQFNPDIYTFPIIVHFGIRPA 780
DB 721 NMLQWEELEMOKAYAECKGMIVTNPGTKPSSVRIIDQIDREQFNPDIYTFPIIVHFGIRPA 780
QY 781 QLSYAGDPQYQKLKMSYVYKLRIHLANS PKVKQTDKOKLAQREBALOKIROKNTMRREVTY 840
DB 781 QLSYAGDPQYQKLKMSYVYKLRIHLANS PKVKQTDKOKLAQREBALOKIROKNTMRREVTY 840
QY 841 ELSSQGFMTGTGIRSDVCOHAMMLPVLTHTIRYHQCLMHLDKLIGYTFQDRCLQLAMTHP 900
DB 841 ELSSQGFMTGTGIRSDVCOHAMMLPVLTHTIRYHQCLMHLDKLIGYTFQDRCLQLAMTHP 900
QY 901 SHHLNFGKMPDHAARNLSNCGIRQPKYGDRAKHHMMRKKGINTLTINISRLQODDPTPS 960
DB 901 SHHLNFGKMPDHAARNLSNCGIRQPKYGDRAKHHMMRKKGINTLTINISRLQODDPTPS 960
QY 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGGLATYRTAIVQNOHMLAMAKKLELD 1020
DB 961 RINHNRELEFLGDVAVVEFLTSVHLVYLPSPLEEGGLATYRTAIVQNOHMLAMAKKLELD 1020
QY 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGA VLEGSLEBAKOLFGRLLFNDPDLREVWLN 1080
DB 1021 FMLYAHGPDLCRESDLRHAMANCFEALIGA VLEGSLEBAKOLFGRLLFNDPDLREVWLN 1080

US-11-001-993-2
; Sequence 2, Application US/11001993
; Publication No. US20050159384A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNAse III And Compositions And Uses Thereof
; FILE REFERENCE: ISI5003-104 (ISI5-5030US.D1)
; CURRENT FILING DATE: 2004-12-02
; PRIOR APPLICATION NUMBER: 10/079,185
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-001-993-2

Query Match 100.0%; Score 7500; DB 6; Length 1374;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MMQGNCHMSFPHGCGCPGSGHGAAPSAPFRPQNLALLHPQPVQYQYEPPSAPS 60
1 MMQGNCHMSFPHGCGCPGSGHGAAPSAPFRPQNLALLHPQPVQYQYEPPSAPS 60
61 TTFSNPAVFLPRPDPVPPPPPPSAQGLPPECIRPPFPNNHQRHPVPPPCPPM 120
61 TTFSNPAVFLPRPDPVPPPPPPSAQGLPPECIRPPFPNNHQRHPVPPPCPPM 120
61 TTFSNPAVFLPRPDPVPPPPPPSAQGLPPECIRPPFPNNHQRHPVPPPCPPM 120
121 PPMPCPNPVPVGAAPGCGTFFPMMPPPSMPPPPPPVPPQVNTQYPPGYSNHPPP 180
121 PPMPCPNPVPVGAAPGCGTFFPMMPPPSMPPPPPPVPPQVNTQYPPGYSNHPPP 180
121 PPMPCPNPVPVGAAPGCGTFFPMMPPPSMPPPPPPVPPQVNTQYPPGYSNHPPP 180
181 SFNSFQNNSSFLPSANNSSPHFRLLPYPLPKASERRSPERLKHVDDHRHDSHGR 240
181 SFNSFQNNSSFLPSANNSSPHFRLLPYPLPKASERRSPERLKHVDDHRHDSHGR 240
241 GERHRLDRRERGRSPDRRQDSRYSDYDGRTPSRHRSYERSRERERHHRDNRSS 300
241 GERHRLDRRERGRSPDRRQDSRYSDYDGRTPSRHRSYERSRERERHHRDNRSS 300
301 PSLERYSKKEYSKSGSYGLSVVPBAGCTPLPGLIKNTDWAAPLEIVNHRSSPREK 360
301 PSLERYSKKEYSKSGSYGLSVVPBAGCTPLPGLIKNTDWAAPLEIVNHRSSPREK 360
301 PSLERYSKKEYSKSGSYGLSVVPBAGCTPLPGLIKNTDWAAPLEIVNHRSSPREK 360
361 KKAARWEEEDKRDNDSSGCKDKYTSIKEKEPEETMPDKNEEELIKPVMIRCHSEN 420
361 KKAARWEEEDKRDNDSSGCKDKYTSIKEKEPEETMPDKNEEELIKPVMIRCHSEN 420
421 YSSDMDVQDSTVGTSLRLDLYKFEELGSRQEKAAAPRPMBPRTKLDLEDLESS 480
421 YSSDMDVQDSTVGTSLRLDLYKFEELGSRQEKAAAPRPMBPRTKLDLEDLESS 480
481 SSSSCSDSDSTSSSSDSSEVPVIAEIKKKAHPRLHDELTYNDPGQNDPCLCKSA 540
481 SSSSCSDSDSTSSSSDSSEVPVIAEIKKKAHPRLHDELTYNDPGQNDPCLCKSA 540
541 KKAARWEEEDKRDNDSSGCKDKYTSIKEKEPEETMPDKNEEELIKPVMIRCHSEN 600
541 KKAARWEEEDKRDNDSSGCKDKYTSIKEKEPEETMPDKNEEELIKPVMIRCHSEN 600
601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEBMPENFCVGLFSLFLRDILELY 660
601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEBMPENFCVGLFSLFLRDILELY 660

DB 601 FEGFSMFAHAPLTNIPLCVIRFNIDYTHFIEBMPENFCVGLFSLFLRDILELY 660
QY 661 DMLAKGPLFEDSPCCPRFFHFMRFVFLPDGKEVLSMHQILLYLRCSKALVPEBEIA 720
DB 661 DMLAKGPLFEDSPCCPRFFHFMRFVFLPDGKEVLSMHQILLYLRCSKALVPEBEIA 720
QY 721 NMLQWEELEWQKAECKGMIVTNPCTKSSVRIIDOLREOFNDVITPPIYHFGIRPA 780
DB 721 NMLQWEELEWQKAECKGMIVTNPCTKSSVRIIDOLREOFNDVITPPIYHFGIRPA 780
QY 781 QLSYADDPQYQKLMKSYVULRHLLANSPKYKOTDOKLQOREALOKIRQKNTMBREVTU 840
DB 781 QLSYADDPQYQKLMKSYVULRHLLANSPKYKOTDOKLQOREALOKIRQKNTMBREVTU 840
QY 841 ELSSQGFWMKTGIRSDVCOHAMMLPVLTHNIRYQCMLHDKLIGYFQDRCLQLAMTHP 900
DB 841 ELSSQGFWMKTGIRSDVCOHAMMLPVLTHNIRYQCMLHDKLIGYFQDRCLQLAMTHP 900
QY 901 SHHNFQAMPDHARNSLNSGIRQPKYGDRCVHMMRRKKGINTLINIMSLGODDPTPS 960
DB 901 SHHNFQAMPDHARNSLNSGIRQPKYGDRCVHMMRRKKGINTLINIMSLGODDPTPS 960
QY 961 RINHNRLERFLGAVVEFLTSVHLTYLPSLEEGGLATYRTALVONOHMLAKKLELD 1020
DB 961 RINHNRLERFLGAVVEFLTSVHLTYLPSLEEGGLATYRTALVONOHMLAKKLELD 1020
QY 1021 FMLYAHGPDLCRESDDRHAMANCFEALIGAVYLEGSLEAKQFLGLFNDPDLREVMNL 1080
DB 1021 FMLYAHGPDLCRESDDRHAMANCFEALIGAVYLEGSLEAKQFLGLFNDPDLREVMNL 1080
QY 1081 YPLHPLQOEPTNDROLIETSPVLQKLEFEBAIGVIFTHVRLABAFILRTVGFNHLTL 1140
DB 1081 YPLHPLQOEPTNDROLIETSPVLQKLEFEBAIGVIFTHVRLABAFILRTVGFNHLTL 1140
QY 1141 GHNQRMFLGDSIMOLVATETYLFIHPDHHGHLTLRSSLVNNRTOAKAEEELGMOEYA 1200
DB 1141 GHNQRMFLGDSIMOLVATETYLFIHPDHHGHLTLRSSLVNNRTOAKAEEELGMOEYA 1200
QY 1201 ITNDKTRPVGLTETKTLADLLESFIALALYTDKOLEVYHFMNVCFFPRLKEFLINDQND 1260
DB 1201 ITNDKTRPVGLTETKTLADLLESFIALALYTDKOLEVYHFMNVCFFPRLKEFLINDQND 1260
QY 1261 PKSQLOCCCLTLTEGKEPDIPLYKTLQTVGSHARTYTVAVYFKBERIGCGKPSIOQA 1320
DB 1261 PKSQLOCCCLTLTEGKEPDIPLYKTLQTVGSHARTYTVAVYFKBERIGCGKPSIOQA 1320
QY 1321 EMGAAMDALSKYFPQMAHQKRFIGKRYROELKEMWREHOREPDETEDIK 1374
DB 1321 EMGAAMDALSKYFPQMAHQKRFIGKRYROELKEMWREHOREPDETEDIK 1374

RESULT 6
US-10-723-860-281
; Sequence 281, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Azitz, Natascha
; APPLICANT: Zlotnick, Wendy M.
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 281
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-281

Query Match 99.8%; Score 7486; DB 5; Length 1374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MMQGNVTCRMSFHPGRCGRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSAPS 60
DB 1 MMQGNVTCRMSFHPGRCGRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSAPS 60
QY 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPLRPPPNHOMRHPFVPCFPFM 120
DB 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPLRPPPNHOMRHPFVPCFPFM 120
QY 121 PPMPCPNPVPVGAAPGCGTFPFMMPPPSMHPPPPVMPOQVNYQYPGYSHNFP 180
DB 121 PPMPCPNPVPVGAAPGCGTFPFMMPPPSMHPPPPVMPOQVNYQYPGYSHNFP 180
QY 181 SFNSFQNNSSFLPSANSSSPHFRHLPYPLPKASERSPERLKHVDHHRDHSGR 240
DB 181 SFNSFQNNSSFLPSANSSSPHFRHLPYPLPKASERSPERLKHVDHHRDHSGR 240
QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTRSRHSYERSRERERHRHNRNS 300
DB 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTRSRHSYERSRERERHRHNRNS 300
QY 301 PSLERSYKKEYSKSGSYGLSVPEPAGCTPELPGELIKNTDSWAPLEIVNHRSPSRK 360
DB 301 PSLERSYKKEYSKSGSYGLSVPEPAGCTPELPGELIKNTDSWAPLEIVNHRSPSRK 360
QY 361 KARWEEKORWSDNOSGSKDNKXYSIKKEPEETMPDKNEEBEELKPVIRCTHSN 420
DB 361 KARWEEKORWSDNOSGSKDNKXYSIKKEPEETMPDKNEEBEELKPVIRCTHSN 420
QY 421 YSSSPMDVGVSTVGTSLRLDLVKFEEELSGRREKAKARPPEPKTKLDEDESS 480
DB 421 YSSSPMDVGVSTVGTSLRLDLVKFEEELSGRREKAKARPPEPKTKLDEDESS 480
QY 481 SESECESDSDSTCSSSDSEVPDVAIEIKRKAHPRLHDELMYNDPGQMDGPLCKCSA 540
DB 481 SESECESDSDSTCSSSDSEVPDVAIEIKRKAHPRLHDELMYNDPGQMDGPLCKCSA 540
QY 541 KARRTGIRHSIYGEBAIKPCRPMTNNAAGLPHYRITVSPPTNFLTDRPTVIEYDHEYI 600
DB 541 KARRTGIRHSIYGEBAIKPCRPMTNNAAGLPHYRITVSPPTNFLTDRPTVIEYDHEYI 600
QY 601 PEGFSMFAAPLNTPLCKVIRFNIDYTHFLEEMMPENPCVAGLEFLPLFRDILEY 660
DB 601 PEGFSMFAAPLNTPLCKVIRFNIDYTHFLEEMMPENPCVAGLEFLPLFRDILEY 660
QY 720 661 DNMLKGPLFEDSPPCPRFHFMRFVRPLPDGKEVLSMHQILLYLRKSKALVPEEEL 720
DB 661 DNMLKGPLFEDSPPCPRFHFMRFVRPLPDGKEVLSMHQILLYLRKSKALVPEEEL 720
QY 780 721 NMLQWELEWQYAECEKGMIVTNGTKDSVRIIDLDBEONPDVITPTIIVHGIRPA 780
DB 721 NMLQWELEWQYAECEKGMIVTNGTKDSVRIIDLDBEONPDVITPTIIVHGIRPA 780
QY 840 781 QLSYADPOYQKLMKSYVLCRLHLANSPKVKQDDKLAQREBALOKIKOKMTREYV 840
DB 781 QLSYADPOYQKLMKSYVLCRLHLANSPKVKQDDKLAQREBALOKIKOKMTREYV 840
QY 900 841 ELSSQGFMTGIRSDVCOHAMMLPVLTTHIRYHQCMLHLDKLGTFFORCLQLMTMP 900
DB 841 ELSSQGFMTGIRSDVCOHAMMLPVLTTHIRYHQCMLHLDKLGTFFORCLQLMTMP 900
QY 960 901 SHHLNGMNPDAHARNSLSNGIRQPYGDRKVVHMMRKKGINTLINISRLGODDPTS 960
DB 901 SHHLNGMNPDAHARNSLSNGIRQPYGDRKVVHMMRKKGINTLINISRLGODDPTS 960
QY 1020 961 RINHRLEFLGDAVVEFLTSVHLVTLFSLEREGGLATYRTALVQNOHMLAKKLELP 1020
DB 961 RINHRLEFLGDAVVEFLTSVHLVTLFSLEREGGLATYRTALVQNOHMLAKKLELP 1020
```

```
QY 1021 FMLVNHGDLGRESPLRHRMANCFEALIGAVYLBESSLEBAQLFGRLLFNDPLREVMLN 1080
DB 1021 FMLVNHGDLGRESPLRHRMANCFEALIGAVYLBESSLEBAQLFGRLLFNDPLREVMLN 1080
QY 1081 YPLHFLQOEPTDQLETSPLVKLTFEPEAIVIFTHVYLLARAFTLRVGFNHLTL 1140
DB 1081 YPLHFLQOEPTDQLETSPLVKLTFEPEAIVIFTHVYLLARAFTLRVGFNHLTL 1140
QY 1141 GHNQMEFLGDSIMQVATEYLFIHPDHHEGLTLSSLVNNTQAKVAEELGMEYA 1200
DB 1141 GHNQMEFLGDSIMQVATEYLFIHPDHHEGLTLSSLVNNTQAKVAEELGMEYA 1200
QY 1201 ITNDKTRPVGLRTTLADLBSFLAAYYTDQLEVTYNTFMVNCFPPLKEETLNQMD 1260
DB 1201 ITNDKTRPVGLRTTLADLBSFLAAYYTDQLEVTYNTFMVNCFPPLKEETLNQMD 1260
QY 1261 PXSQLOCCCLTLRTGKSPDIPLYKTLQTVGSHARTYVAVYFGERIGCGSPSIQQA 1320
DB 1261 PXSQLOCCCLTLRTGKSPDIPLYKTLQTVGSHARTYVAVYFGERIGCGSPSIQQA 1320
QY 1321 EMGAAMDALFKYNPOMAHQKRFIGRKYROELKERNWEREHQERBPDETDIKK 1374
DB 1321 EMGAAMDALFKYNPOMAHQKRFIGRKYROELKERNWEREHQERBPDETDIKK 1374
```

```
RESULT 7
US-10-756-149-4809
; Sequence 4809, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Naeem
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patencin version 3.2
; SEQ ID NO 4809
; LENGTH: 1374
; TYPE: PRK
; ORGANISM: Homo Sapiens
US-10-756-149-4809
```

Query Match 99.8%; Score 7486; DB 5; Length 1374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MMQGNVTCRMSFHPGRCGRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSAPS 60
DB 1 MMQGNVTCRMSFHPGRCGRGCGHAGARPASFRPQNLRLHPQOPVQOYEPPSAPS 60
QY 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPLRPPPNHOMRHPFVPCFPFM 120
DB 61 TTFSSNPAENFLPRPDPVFPFPPMPPSAQGLPCPLRPPPNHOMRHPFVPCFPFM 120
QY 121 PPMPCPNPVPVGAAPGCGTFPFMMPPPSMHPPPPVMPOQVNYQYPGYSHNFP 180
DB 121 PPMPCPNPVPVGAAPGCGTFPFMMPPPSMHPPPPVMPOQVNYQYPGYSHNFP 180
QY 181 SFNSFQNNSSFLPSANSSSPHFRHLPYPLPKASERSPERLKHVDHHRDHSGR 240
DB 181 SFNSFQNNSSFLPSANSSSPHFRHLPYPLPKASERSPERLKHVDHHRDHSGR 240
QY 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTRSRHSYERSRERERHRHNRNS 300
DB 241 GERHSLDRERGRSPDRRQDSRYSDYDRGTRSRHSYERSRERERHRHNRNS 300
QY 301 PSLERSYKKEYSKSGSYGLSVPEPAGCTPELPGELIKNTDSWAPLEIVNHRSPSRK 360
DB 301 PSLERSYKKEYSKSGSYGLSVPEPAGCTPELPGELIKNTDSWAPLEIVNHRSPSRK 360
```

```

QY 361 KARMEEEKDRMSDNOSSGKDKNYTISKKEPEETMPDKNEBEEELKRWIRCHTSHN 420
DB 361 KARMEEEKDRMSDNOSSGDKNYTISKKEPEETMPDKNEBEEELKRWIRCHTSHN 420
QY 421 YSSDPMDOVGSTVGTSTRLNLYKFEPEELGSRBEKAKAARPMPEPRTKLDDELSS 480
DB 421 YSSDPMDOVGSTVGTSTRLNLYKFEPEELGSRBEKAKAARPMPEPRTKLDDELSS 480
QY 481 SESECESDSTCSSSSSEVPFVIAEIKRKAHPRLDELMTYNNPGQNDPGLCKCSA 540
DB 481 SESECESDSTCSSSSSEVPFVIAEIKRKAHPRLDELMTYNNPGQNDPGLCKCSA 540
QY 541 KARRTGIRHSIYGEBAIKPCREPMNAGRLPHYRTVSPNTFLDRPTVIEYDDHEXI 600
DB 541 KARRTGIRHSIYGEBAIKPCREPMNAGRLPHYRTVSPNTFLDRPTVIEYDDHEXI 600
QY 601 PEGFSFPAHAPLNTIPLCKVIRFNIDYTHFIEBMAPEFVCVGLFSLFPRDLIELY 660
DB 601 PEGFSFPAHAPLNTIPLCKVIRFNIDYTHFIEBMAPEFVCVGLFSLFPRDLIELY 660
QY 661 DNMLKGPLFEDSPCCPRFHPMPRFVPLPDGKEVLSHMOILLYLKCSKALVPEEETA 720
DB 661 DNMLKGPLFEDSPCCPRFHPMPRFVPLPDGKEVLSHMOILLYLKCSKALVPEEETA 720
QY 721 NMLQWEELEMOXYAECECKGMIYVTPGTKPSVRIIDLDREQFNPDVITPEIIVHFGIRPA 780
DB 721 NMLQWEELEMOXYAECECKGMIYVTPGTKPSVRIIDLDREQFNPDVITPEIIVHFGIRPA 780
QY 781 QLSYAGDPYOYOKLWMSYVTKLRHLANSPYKQOTDKQLAQREBALOKIRQKNTMREBTV 840
DB 781 QLSYAGDPYOYOKLWMSYVTKLRHLANSPYKQOTDKQLAQREBALOKIRQKNTMREBTV 840
QY 841 ELSSSGFMKTGTJRSVDCOHAHMLPVLTTHHRYHQCLMHLDKLIGYTFORCLQLAMTHP 900
DB 841 ELSSSGFMKTGTJRSVDCOHAHMLPVLTTHHRYHQCLMHLDKLIGYTFORCLQLAMTHP 900
QY 901 SHHLNFGNMPDHAHNSLSNCGIRQPKYGDRCVHMHMKKKGINTLTINISRLQODDPTPS 960
DB 901 SHHLNFGNMPDHAHNSLSNCGIRQPKYGDRCVHMHMKKKGINTLTINISRLQODDPTPS 960
QY 961 RINNNERLEFLGDVAVEFLTSVHLVYLPFSLSEEGIATYRTAIVONQHLAMLAKKLEDP 1020
DB 961 RINNNERLEFLGDVAVEFLTSVHLVYLPFSLSEEGIATYRTAIVONQHLAMLAKKLEDP 1020
QY 1021 FMLYAHGPRLCRESDLRHAMANCEFALIGAVYLEGSLSEAKOLFGLFNDPDLREVMIN 1080
DB 1021 FMLYAHGPRLCRESDLRHAMANCEFALIGAVYLEGSLSEAKOLFGLFNDPDLREVMIN 1080
QY 1081 YPLHPLQLOEPNTDROLIETSPVLQCLTFEEBAIGVIFTHVRLAARFLLRTVGFNHLTL 1140
DB 1081 YPLHPLQLOEPNTDROLIETSPVLQCLTFEEBAIGVIFTHVRLAARFLLRTVGFNHLTL 1140
QY 1141 GHNORMEFLGDSIMQVATEYLFIHPDDHBEGLTLRSLSLVNRTQAVAEELGQEXYA 1200
DB 1141 GHNORMEFLGDSIMQVATEYLFIHPDDHBEGLTLRSLSLVNRTQAVAEELGQEXYA 1200
QY 1201 IINDTKRVRGLRTKTLADLLESFIAALYTDKOLEVHTPMVCFPRPKLEFLINDAMD 1260
DB 1201 IINDTKRVRGLRTKTLADLLESFIAALYTDKOLEVHTPMVCFPRPKLEFLINDAMD 1260
QY 1261 PXSLOQCCLTJRTBEKEPDIPLKYTLQTVGSPHATTVAVYFKERIGCGKSPSIOQA 1320
DB 1261 PXSLOQCCLTJRTBEKEPDIPLKYTLQTVGSPHATTVAVYFKERIGCGKSPSIOQA 1320
QY 1321 EMGAAMDALEKTNFPMAHQKRFIRKYNQELKEMWEROHEREDDETEDIK 1374
DB 1321 EMGAAMDALEKTNFPMAHQKRFIRKYNQELKEMWEROHEREDDETEDIK 1374

```

RESULT 8
 US-11-097-143-2409
 ; Sequence 2409, Application US/11097143
 ; Publication No. US20050208558A1

```

; GENERAL INFORMATION:
; Applicant: Venter, J. Craig
; Applicant: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAY, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2409
; LENGTH: 1327
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-2409

Query Match 39.5%; Score 2963; DB 6; Length 1327;
Best Local Similarity 45.6%; Pred. No. 7,8e-173;
Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;

QY 90 QGPLPCCPIRPPFPNQHMHPPVPPCPMPMPMPCPN-NPPVPGAP-----P 137
DB 3 QPPLPPPVQPA-----PPPPPEEEDLSPPGVSPHMYNSNSHSG 46

QY 138 GQCTPFPMPRPSMPPPPPPVMPPOVNYQYPPGY-----SHHNPP 179
DB 47 SSKSLDYVPEPAPYASS---VPSYDPYQPPAYGEGYAVNEQAOKYGGQESHYQY 103

QY 180 PEPNFSQNNPSSFLPSANSSSPHFNHLPPLPYL---PAPEBRRSPBKLKYDHRND 235
DB 104 PA-----SGSSFLYES---YKYPDRYPAYSSNYPBPSEKQ-----RYTS 139

QY 236 HSHGGERHRSILDRRERGRSPDRRQDSRYRSDYDGRTPSHRSYEREREERERHR 295
DB 140 NSSSQGYHHP-----GYSSGR---RYEQRHDD---EHRQIQDSRYAHEPHNGY 183

QY 296 DNRSPSLERSYKKEYKSGRSGYGLSVPEPACTPELPGELIKNTDSVAPLEIVNHS 355
DB 184 AHRQAKGSHGY---YGSAAH-----QVSDVSPGCHHERERN 219

QY 356 PSREKKRA--RWEEKDR---WSDNOSGKDKNTSIEKEPEETMPDKNEBEEELK 409
DB 220 ETLKTRAKPKYTEERDRLRQMSNFC-----EKPEEDVYKKNALSEADAV 267

QY 410 PVIWIRCHTSHENYSSDPMQVGDSTVYGTSTRLNLYKFEPEELGSRBEKAKAARPMWP 469
DB 268 ESWVRSFPAELYERTK---SENEVGRARQKLTFLFDELLQRAKAVREGLPYVFP 323

QY 470 KTKL-----DELESSESSECEDESDSTCSSSDSEVFDV-----IAEIKKKAH 514
DB 324 PKARARVCCKHKGSEACSSSSSDDDDEDA-----FKIQDCMEELSRKVOH 373

QY 515 PDLHDELMTYNDPGQNDPGLCKSAKARTGIRHSIYGEBAIKPCREPMNAGRLPHY 574
DB 374 PORVHADLMTNDAGEMNDPGLCRCSAKSRIRIGIRGYGETGYKLCDBNSNAGLKFY 433

```

```

QY 575 RTTSGPPTNFTLDRPVIVIEYDDHEIYFEGSGMFAAPLTNIDLCYIRINIDYTHIFIRE 634
DB 434 RISISPTNFTLTKPTIIRKIDHEHLEBGSFSLSHVRLSDLPVCXVIRINIEYTHEEBS 493
QY 635 KMPENCVAGLELFSFLPRDILEYDMNLKGP-----BEDSPCCPRPHMFRFRFP 690
DB 494 KMPENTTIELDFFKYLHELELVDYDMLPVBGNVEBS---CPAHFFPRFRDLP 550
QY 691 DSGKEVLSMHOILLYLLRCSKALVPBEELANMLQWEELEWOKYAECKMITYNPGTKS 750
DB 551 DNGKEVLAVAEVLYRLDNSAQVERQOLHLNQISQSEMONVVDPIKGMALYTKPGYKPC 610
QY 751 SVRIDLDDEQRPNV-----ITPRIVHFGIRPAQLSYAGDPQYOKLMKSVYK 801
DB 611 STRVOLDLN--NSDLPECVDRBTGISHPAIVHFGICHPQLSYAGNPEYQKAMREYKXR 668
QY 802 HLLANPKYQTDKOKLAOREEALOKIROKNTRRREVTVELSOGWKGISDVCOAHM 861
DB 669 HLMANMSKSPKDKRKLBEKEOLOEMKTQGRMKRITVAISSEGYRTGIMCDVYQAHM 728
QY 862 MDPVLTNHIRYHOCLMHLMDKLIGYTFQDBCLQLAMTPSHHLNFGMNPDHANSLSCNG 921
DB 729 LIPVLGLHLPFKSLDLEESIGYRPNRYLQLALTHPSYKENVGTNPDHANSLSCNG 788
QY 922 IQPKYGDKNVHMHRKKGINTLINMSRLGDDPTPSRINNELELELGDVAVELTS 981
DB 789 IQOPEYGDKNVHMHRKKGINTLINMSRLGDDPTPSRINNELELELGDVAVELTS 848
QY 982 VHLVYFSPLESGGLATYRTAIVONOHMLAKKLELDPMLYAHGPDLCRESDLRHANA 1041
DB 849 IHLFEPFLESGGLATYRTAIVONOHMLAKKLELDPMLYAHGPDLCHELELHANA 908
QY 1042 NCFEALIGAVYLEGSLEAKOLFGRLLF--NDPDLREVMNLYPLHLOQEPNTRQLIET 1100
DB 909 NCFEALMGALLDGGIKVADVEFTDALPQDEKLISIMKNLPEHPLQEBEPIDGRCIDS 968
QY 1101 SPVLOKLTPEBEAIGYITFHVRLAARFRLTYGFMHLTLGNORMEFIDGSIQOLVATE 1160
DB 969 YVLLKELTFEBSIGIKFKHIRILARAFDRSIGFHLTLGSGOBLFLEFIDVYLOLCSB 1028
QY 1161 YFIFHRDHBEGLTLRSLVNNRFOAKYAEELGMOEYAL--TNDTKRPVGRTRTTLAD 1219
DB 1029 YFIRHPEHBEGLTLRSLVNNRFOAKYAEELGMOEYAL--ADLKYTORAD 1084
QY 1220 LLESFIALYTDKLEEVHTFMNVCFEPRLKEFILLQDNMDPKSQLOQCLTLRT--BGKE 1278
DB 1085 LLEAFIAGALYVDKGLYCFOFCHVCLFPHLOQLFVNMQDNMDPKSKLOQCLTLRTMDGB 1144
QY 1279 PDIPLYKLTQTVGSHARTYTAIVYFKGERIGCGKPSIQOAEKMAAMALEKTN--FPQ 1336
DB 1145 PDIPYKVVVEASGPTVTRVYKVAIVYFRSKRLATSSGSSIQOAEKMAAKALENSRDLFPQ 1204
QY 1337 MAHOKRFIGRKYROEL---KEMWREHOREPDEDEDIK 1374
DB 1205 LDHOKRVIAMSKIGOTGNELDNDSDROHE-----EKIKR 1239

```

RESULT 9
US-10-205-331-116

```

/ Sequence 116, Application US/1020531
/ Publication No. US20040058326A1
/ GENERAL INFORMATION:
/ APPLICANT: Warner-Lambert Company
/ APPLICANT: Lee, Kevin
/ APPLICANT: Dixon, Alister
/ APPLICANT: Brooksbank, Robert
/ APPLICANT: Finnock, Robert
/ TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
/ FILE REFERENCE: WL-A-018199
/ CURRENT APPLICATION NUMBER: US/10/205,331
/ PRIOR FILING DATE: 2002-07-24
/ PRIOR APPLICATION NUMBER: GB 0118354.0
/ PRIOR FILING DATE: 2001-07-27

```

```

/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 116
/ LENGTH: 541
/ TYPE: PRF
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Putative ribonuclease III
US-10-205-331-116

```

```

Query Match 38.0%; Score 2847; DB 4; Length 541;
Best Local Similarity 99.3%; Pred. No. 3,2e-166;
Matches 537; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 834 MREYTVLSGSGFWKTGIRSDVCOAHMLPVLTNHIRYHOCLMHLMDKLIGTFQDBCLL 893
DB 1 MREYTVLSGSGFWKTGIRSDVCOAHMLPVLTNHIRYHOCLMHLMDKLIGTFQDBCLL 60
QY 894 OLAMTHPSHHLNFGMNPDHANSLSCNGIROPKYGDKNVHMHRKKGINTLINMSRLG 953
DB 61 OLAMTHPSHHLNFGMNPDHANSLSCNGIROPKYGDKNVHMHRKKGINTLINMSRLG 120
QY 954 ODDPTPSRINNELELELGDVAVELTSVHLYLPPSLEGGATYRTAIVONOHMLA 1013
DB 121 QDDPTPSRINNELELELGDVAVELTSVHLYLPPSLEGGATYRTAIVONOHMLA 180
QY 1014 KKLIEDPMLYAHGPDLCRESDLRHANMCEPALIGAVYLBGSLEBAKOLFGRLLFPNDP 1073
DB 181 KKLIEDPMLYAHGPDLCRESDLRHANMCEPALIGAVYLBGSLEBAKOLFGRLLFPNDP 240
QY 1074 LREVMNLYPLHLOQEPNTRQLIETSPVLOKLTPEBEAIGYITFHVRLAARFRLTY 1133
DB 241 LREVMNLYPLHLOQEPNTRQLIETSPVLOKLTPEBEAIGYITFHVRLAARFRLTY 300
QY 1134 GFNHLTLGNORMEFIDGSIQOLVATEYLFIFPDHBEGLTLRSLVNNRTOAKYAE 1193
DB 301 GFNHLTLGNORMEFIDGSIQOLVATEYLFIFPDHBEGLTLRSLVNNRTOAKYAE 360
QY 1194 LGMOEYAITNDTKRPVGRTRTTLADLLESFIALYTDKLEEVHTFMNVCFEPRLKEP 1253
DB 361 LGMOEYAITNDTKRPVGRTRTTLADLLESFIALYTDKLEEVHTFMNVCFEPRLKEP 420
QY 1254 LNDMDPKSQLOQCLTLRTGKBPDIPLYKLTQTVGSHARTYTAIVYFKGERIGCGK 1313
DB 421 LNDMDPKSQLOQCLTLRTGKBPDIPLYKLTQTVGSHARTYTAIVYFKGERIGCGK 480
QY 1314 GPSIOAEKMAAMALEKTNFPQMAHOKRFIGRKYROELKEMWREHOREPDEDEDIK 1373
DB 481 GPSIOAEKMAAMALEKTNFPQMAHOKRFIGRKYROELKEMWREHOREPDEDEDIK 540
QY 1374 K 1374
DB 541 K 541

```

RESULT 10
US-10-774-974-37

```

/ Sequence 37, Application US/1074974
/ Publication No. US20040126867A1
/ GENERAL INFORMATION:
/ APPLICANT: Wu, Hongjiang
/ APPLICANT: Crooke, Stanley T.
/ TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
/ FILE REFERENCE: ISIS029/ISPH-0522
/ CURRENT APPLICATION NUMBER: US/10/774,974
/ PRIOR FILING DATE: 2004-02-09
/ PRIOR APPLICATION NUMBER: US/09/900,425B
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 37
/ LENGTH: 466
/ TYPE: PRF

```



```
; ORGANISM: Homo sapiens
US-10-774-974-37

Query Match      32.8%; Score 2463; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 9.9e-143;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 909 NPDHANSISNGCIGRPKYGDRKYHHMHRKKGINTLINIMSRIGODDPPSRININERL 968
DB 1 NPDHANSISNGCIGRPKYGDRKYHHMHRKKGINTLINIMSRIGODDPPSRININERL 60
QY 969 EFLGDAVVEFLTSVHLVYLPPSLIEEGGLATYRTAIVONOHMLAKKLEIDPMLYAHGP 1028
DB 61 EFLGDAVVEFLTSVHLVYLPPSLIEEGGLATYRTAIVONOHMLAKKLEIDPMLYAHGP 120
QY 1029 DICRESDDLHMANANCEALIGAYYLEGSLIEEAKQLFNDPDLREVWNLPLHPLDL 1088
DB 121 DICRESDDLHMANANCEALIGAYYLEGSLIEEAKQLFNDPDLREVWNLPLHPLDL 180
QY 1089 QEPNTDROLIETSPVLQKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHNORMEF 1148
DB 181 QEPNTDROLIETSPVLQKTEFEBAIGVIFTHVRLARAFTLRTVGFNHLTLGHNORMEF 240
QY 1149 LGDSIMQLVATEYLFHPDHEGHLTLRSSLVNNRTQAKVAELGMOEYAITNDKTKR 1208
DB 241 LGDSIMQLVATEYLFHPDHEGHLTLRSSLVNNRTQAKVAELGMOEYAITNDKTKR 300
QY 1209 PVGLRKTLDLLESTIALYTDKLEFYHTFMNVCFPRPKKEFLINQDMDPKSQLOQC 1268
DB 301 PVGLRKTLDLLESTIALYTDKLEFYHTFMNVCFPRPKKEFLINQDMDPKSQLOQC 360
QY 1269 CILTRTEGKEPDIPLKTLQTVGSHARTYTVAVYFKGERIGCGKPSIQOAEAGAMDA 1328
DB 361 CILTRTEGKEPDIPLKTLQTVGSHARTYTVAVYFKGERIGCGKPSIQOAEAGAMDA 420
QY 1329 LEKYNFPQMAHQKRFGRKYRQELKEMREREHQEREPDETEDIKK 1374
DB 421 LEKYNFPQMAHQKRFGRKYRQELKEMREREHQEREPDETEDIKK 466

RESULT 11
US-10-103-313-471
; Sequence 471, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 471
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-471

Query Match      26.1%; Score 1961; DB 4; Length 378;
Best Local Similarity 98.9%; Pred. No. 4.9e-112;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 997 ATYRTAIVONOHMLAKKLEIDPMLYAHGPDLCRESDDLHMANANCEALIGAYYLEGS 1056
DB 1 ATYRTAIVONOHMLAKKLEIDPMLYAHGPDLCRESDDLHMANANCEALIGAYYLEGS 60
QY 1057 LBEAKQLFGRLLFNDDLEREVWNLPLHPLQLOEPTDRLIETSPVLQKTEFEBAIGV 1116
DB 61 LBEAKQLFGRLLFNDDLEREVWNLPLHPLQLOEPTDRLIETSPVLQKTEFEBAIGV 120
QY 1117 IFTHVRLARAFTLRTVGFNHLTLGHNORMEF LGDSIMQLVATEYLFHPDHEGHLTL 1176
```

```
DB 121 IFTHVRLARAFTLRTVGFNHLTLGHNORMEF LGDSIMQLVATEYLFHPDHEGHLTL 180
QY 1177 LRSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRKTLDLLESTIALYTDKLEFY 1236
DB 181 LRSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRKTLDLLESTIALYTDKLEFY 240
QY 1237 VHTFMNVCFPRPKKEFLINQDMDPKSQLOQCCLTLRTGKEPDIPLKTLQTVGSHAR 1296
DB 241 VHTFMNVCFPRPKKEFLINQDMDPKSQLOQCCLTLRTGKEPDIPLKTLQTVGSHAR 300
QY 1297 TYTVAVYFKGERIGCGKPSIQOAEAGAMDALIEKNFPQMAHQKRFGRKYRQELKEMR 1356
DB 301 TYTVAVYFKGERIGCGKPSIQOAEAGAMDALIEKNFPQMAHQKRFGRKYRQELKEMR 360
QY 1357 WEREHQEREPDETEDIKK 1374
DB 361 WEREHQEREPDETEDIKK 378

RESULT 12
US-10-103-313-307
; Sequence 307, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 307
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-313-307

Query Match      17.6%; Score 1320; DB 4; Length 263;
Best Local Similarity 97.7%; Pred. No. 6.7e-73;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1118 FTHVRLARAFTLRTVGFNHLTLGHNORMEF LGDSIMQLVATEYLFHPDHEGHLTL 1177
DB 7 FTHVRLARAFTLRTVGFNHLTLGHNORMEF LGDSIMQLVATEYLFHPDHEGHLTL 66
QY 1178 RSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRKTLDLLESTIALYTDKLEFY 1237
DB 67 RSSLVNNRTQAKVAELGMOEYAITNDKTKRPVGLRKTLDLLESTIALYTDKLEFY 126
QY 1238 HTFMNVCFPRPKKEFLINQDMDPKSQLOQCCLTLRTGKEPDIPLKTLQTVGSHART 1297
DB 127 HTFMNVCFPRPKKEFLINQDMDPKSQLOQCCLTLRTGKEPDIPLKTLQTVGSHART 186
QY 1298 TYTVAVYFKGERIGCGKPSIQOAEAGAMDALIEKNFPQMAHQKRFGRKYRQELKEMR 1357
DB 187 TYTVAVYFKGERIGCGKPSIQOAEAGAMDALIEKNFPQMAHQKRFGRKYRQELKEMR 246
QY 1358 WEREHQEREPDETEDIKK 1374
DB 247 WEREHQEREPDETEDIKK 263

RESULT 13
US-09-900-425A-3
; Sequence 3, Application US/09900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongliang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNAse III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900,425A
```



```

; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 412
; TYPE: prt
; ORGANISM: Caenorhabditis elegans
US-09-900-425A-3

```

Query Match	11.9%	Score 890;	DB 3;	Length 412;
Best Local Similarity	43.1%	Pred. No. 3e-46;		
Matches 181;	Conservative	85;	Mismatches 136;	Indels 18;
				Gaps 77;
QY	944	TLINIMSRIGODDPTPSRININNEERLEFGDAVEEFTSVLYYLFPSEEGGLATYRTAI	1003	
Db	2	SLFPMIKOTSGSEP----ILHNERLEYIGDAVEELIVSHHLYFMLTHHFEGLATYRTAL	57	
QY	1004	VONCHLMTLAKKLELDPFMYLAHGBDLCREDDLRHAMNCEBALGAYLEGSLEEAQOL	1065	
Db	58	VONENLMTLAKNCRIDEMLOQSHGADLLNVAEFKRLANAEAVDAALYLDGGALPCVYI	117	
QY	1064	FGRLLE-NDPDLREVLWNLVPLPLDLOEPNDPDRQLEIETSPYLOKLTBEBAAGVIFTVR	1122	
Db	118	FSKAMYGHQPLKKEKMDHINHEHLKREDPGQGRDLSFTPLTSTFHALEERLGIOFNIR	177	
QY	1123	LLAAPFLRTVGFNHLTGNORMFLEGDSTIMQVATYELTPIHPDHHGHILTLRSLV	1182	
Db	178	LLAAPFTRFRNLPNNDLTGHNORLEWLDSDVQLQVDSFLRFRFPYHGHMSILRTSLV	237	
QY	1183	NNRTQAKAEELIGMOEYAIITNDKTRPV---GLRKTTLADLSEFIALYDVKOLEYHT	1239	
Db	238	SNQDQAVVCDLDGLTEFYI-----KAPKYTELKIKDKADIVLEATIGLYVDYRGHEHRA	292	
QY	1240	FMANVCFPRLKEFIINQMDNPKSOLQOCCLTLR-TESEKPDIPLYKTLQTVGSPHARTY	1298	
Db	293	FIRIVFCERLKGFIETSEKMDAKSHLQOQCLAMRDPSSSEDPMEYRYLTGIEGPTNNRIF	352	
QY	1299	TVAYVEKGERIGCGGSPSIIQOAEKMA---AMDALETKYNPOM-ANQKRFIQRKYRQELKE	1354	
Db	353	KIAYVYKRRILASAEVNHKELVLAETLALNTLESMSFSKKAKANSNMRRRLQEDTSD	412	

```

RESULT 14
US-10-079-185--3
; Sequence 3, Application US/10079185
; Publication No. US2003044941A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III And Compositions And Uses Thereof
; FILE REFERENCE: 1S1S030
; CURRENT APPLICATION NUMBER: US/10/079,185
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 09/479,783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 08/659,440
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: 09/900,425
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-079-185--3

```

	Query Match	Best Local Similarity	11.9%;	Score 890;	DB 4;	Length 412;
	Matches 181;	Conservative 85;	Mismatches 136;	Indels 18;	Gaps 7;	
Qy	944	TLINISRLQDDPBTBRIRINHNRLERLEGRVAVVELTIVNLYLFLFSELEGGATATTTAT	1003			
Db	2	SLFIMKGTISGGGP-----IHNRLERLEGRVAVVELIYNHLYPMLTNNFEGGATATTTAT	57			
Qy	1004	VONOHMLAKKLELDLPMLYAGRPDLCRESDBLRHMANCFEALIGAVULEGSLEEKOL	1063			
Db	58	VONENLTLTKNGRIDEMLOYSHGADILINVAERKHALANFEAVMAIYLDGGLAPCDVI	117			
Qy	1064	FGRLLE-NDEDLREVMINYLPLHLPLQEPNTDROLIETSVLOKLTFFEEAIGVITFNR	1122			
Db	118	FSKAMYGHOVLKEXKMDHINEHELKREDPQGDRLDLSFTPTLSTFHALERLGIQFNIR	177			
Qy	1123	LIRAFPLTRVGNRNHLTLGNHOMERFLGDSIMOLVATEVLEFINRRDHNEHLLTLSSV	1182			
Db	178	LIRAFPLFRNRIPNNNDLTGKGNORLEWIGDSVLQIVSDPLFYRRPYNHEGMSLRTSLV	237			
Qy	1183	NNRTQAKVAEELGQOEYAITNDTKLPV---GLRTKLADLLSEFIALYTDKDLLEVHT	1239			
Db	238	SNQGVAVVCCDLLEPTEFVI-----KAPYKTRPELTKOKADLVAFIALVVDREIHCRA	292			
Qy	1240	FMNVCFEPLRKEFLINDANDPYSQLOOCCLTLR-TEGKEDPIPLYLTLQTVGSHARTY	1298			
Db	293	FIRIVCFEPLRKHFIETSEKMDAKSHILOQWCLAMRDPSSESDMPEYVLVIGETPNRIIF	352			
Qy	1299	TVNVYFCGERIIGGKGPSIQOAEWGA---AMDLEKYNPFOM-AHQRFIGRKYQRLKE	1354			
Db	353	KIATVYGGKLAGAESNVHKAELRVALLADLNEHSFESKMAKXNNNRRLREQDTS	412			

[illegible]

THIS PAGE BLANK (USPTO)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)


```
Db 579 RS-----RSRSPRYRRYHSRSSPQAFRCQSPKKNVPOGETEREYFNRYREVPPEYDMK 633
Qy 290 -----ERRHRDRNRSPSLERSYKYGSGSYGLSVPEPACETELPGE 336
Db 634 AYVGSVDPRDPPEKERYE-----WEKRYEWYKYYKGAAGQPPPSA----- 679
Qy 337 IIKNTDSMAP-----PLEYVNHNSPSREKKRARWEEKRWSNDQSG-----KD 381
Db 680 ---NRNFSPEPRLPLINRN---SPFTRGRREDYVGCQSHRSRNTGSGNPEKLSARDGHNQ 734
Qy 382 KNYTSLKEKEPEETMD-----KNEESELKPKWIRCTHSENYSSDPM--- 427
Db 735 KONTSKESKESEVAPDGGKGNKHKRKRKKESESGFLNPELLETSRK---SREPTGV 790
Qy 428 -DQVGDSTVVGTSR-----LRDLYDKFEE---LGSROKAVAAPPEPPTKLEDELE 478
Db 791 EBNKTDLSLVLSRDATAVRD--EPMDBESTIPKSVSEKDKRER---DKPAKGDKTKR 845
Qy 479 SSSESECESEDSTCSSSDSE-----VPDVIAEIRKKAHPDRL 518
Db 846 KNDGSVSKKENIVKPAKGPOEKVDGDVRLDLNLQLKKPKRRLRL 893
```

RESULT 2

```
US-10-793-626-1658
; Sequence 1658, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMELLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1658
; LENGTH: 245
; TYPE: PRM
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1658
```

```
Query Match 3.5%; Score 260.5; DB 6; Length 245;
Best Local Similarity 30.4%; Pred. No. 7,2e-10;
Matches 70; Conservative 57; Mismatches 88; Indels 15; Gaps 8;
```

```
Qy 1105 OKLTFEEAIGVIFTHVRLARAFTRLTV--GFNHLTLGHNMREFLGDSSIMQVATERYL 1162
Db 15 OKFTDMKSLGLRFRKIIDLQQAFFSHSPTNDPMNRLHNNELBELGLDAVELTYSRL 74
Qy 1163 FTHPDRHGHLLTLRLSLVNNRTQAKVAELGMQEVATINDTKRPVGLRTK--TLADL 1220
Db 75 FDRHPLRPGNLTKMAATVCEPSLVTFANKIKLNLILLG--KGEKGTGRTTPPSLSDA 133
Qy 1221 LBSFIALYTDKDLKLVNTHPMVNCPPRLK--EFLINOMNDPKSQLOCCCLTLRTGEX 1278
Db 134 FEFVFLALYDQDLSDVMTFAEKVIIPPVEEDDLVGVPF---KTQFQ---YVHSQNK- 186
Qy 1279 PDIPLKTLQTVGPSHARTYTVAVVPKGERIGCGKSPSIQQAEMGAMDA 1328
Db 187 GDV--TYQLKEBPAHRLFTSEVIIENKAVAEKGKTKKESQKAAEQ 235
```

RESULT 3

```
US-11-150-845-40
; Sequence 40, Application US/11150845
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokineics, Inc.
```

```
; APPLICANT: Tomasevici, Nenad
; APPLICANT: Uia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-007720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 40
; LENGTH: 498
; TYPE: PRM
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(498)
; OTHER INFORMATION: SCAR2/WAVE2
US-11-150-845-40
```

```
Query Match 3.0%; Score 224; DB 7; Length 498;
Best Local Similarity 29.5%; Pred. No. 3.7e-07;
Matches 76; Conservative 17; Mismatches 79; Indels 86; Gaps 16;
```

```
Qy 19 PRGRGHARPARPAPRQNLRLHLPQRPVQYQY-----BPP 56
Db 248 PPOQSDASSPS--PFSSEDLN-----PPPAFSPVDNQSGGLAGPKSSVSPSHPP 301
Qy 57 SAPSTFNSPAPNLLPRPDPVPPPPMPSPSAQG--PLRCPD---RPPFPHOMKHPP 112
Db 302 PAPPLGSPGPKKGPAP-----PAPPPPPPMIGIPPPPPVVGSGPCTP-----PP 350
Qy 113 VPCCPPMP---PPMCCNNPVPQAPPGQCTPPPMAPP--SMP---HPPPPVMPQOV 164
Db 351 SPPSFPDPDFAAPP---PPPPA-----DYPTLPPPLSGOPTGAPPPPPPP 398
Qy 165 NYQVPGYSHNPPRPPSPFSPNMPSSFLPSANNSSPHFRLPPYPLKASER----- 219
Db 399 ----PPGP-----PPPTGADGQPAIPPLSDTYK-----PXSPLPVSADANSDLIS 442
Qy 220 --RSPERLKHYYDHRRD 235
Db 443 AIRQGFQRLRVBEGREGE 460
```

RESULT 4

```
US-11-149-945-4
; Sequence 4, Application US/11149945
; Publication No. US20060014266A1
; GENERAL INFORMATION:
; APPLICANT: Cytokineics, Inc.
; APPLICANT: Tomasevici, Nenad
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: PURIFICATION OF ARP2/3 COMPLEX AND COMPOSITIONS CONTAINING
; FILE REFERENCE: 020552-007810US
; CURRENT APPLICATION NUMBER: US/11/149,945
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/578,969
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 498
; TYPE: PRM
; ORGANISM: Homo sapiens
US-11-149-945-4
```



```
Query Match 3.0%; Score 224; DB 7; Length 498;
Best Local Similarity 29.5%; Pred. No. 3.7e-07;
Matches 76; Conservative 17; Mismatches 79; Indels 86; Gaps 16;

Oy 19 PRGRGHGARRPSPRRPONTLRLLHQPPVQYQY-----EPP 56
Db 248 PPPQSDSASSPS-PSFSSEDL-----PPPAEFSYVDNQSGGLAGPKRSSVVSHP 301
Oy 57 SAPSTSSNPAANFPRPDVFPFPPMPPSAQ--PLPPCP--RPPPNQMHPP 112
Db 302 PAPPLGSPGPKGFAP-----PAPPPPPPMIGIPPPPPVGFSGPGR-----PP 350
Oy 113 VPPCPFPPM-----PPMPCPNPPVPCAPGCGTFPPMPP--SMP--HPPPPVMPQOV 164
Db 351 SPPSPFPPHDPFAPPP-----PPPPA-----DYPLPPPLSQPFGAAPPPPPP----- 398
Oy 165 NYOYPRGCGHNNPPRPSFQNNPSSFLPSANSSPHRLPPPLPRAEER----- 219
Db 399 ----PPGP-----PPPPFGADGQPAIPPLSDTK-----PKSSLPAVSDARSDDL 442
Oy 220 --RSPRLKHVDHRRD 235
Db 443 AIRGCFQLRVEREQE 460

RESULT 5
US-11-074-176-136
; Sequence 136, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kjaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-136

Query Match 3.0%; Score 221.5; DB 7; Length 228;
Best Local Similarity 29.0%; Pred. No. 2e-07;
Matches 67; Conservative 47; Mismatches 102; Indels 15; Gaps 6;

Oy 1107 LTFEFAELVIFTHVLLAAFTLRVGFNHLTLG--HNQRMFLDSDIMQVLATEYLF 1164
Db 7 LNKLNKENVIERNEGLLEAFTHSYSNEHPRDDGIRDYKLEFLDVALELVANSYIVR 66
Oy 1165 HPPDHHEGLTLIRSLVNNRTOAKVAELGMO--EYAITNDKTKRPVGLRTKTLADLLS 1223
Db 67 HYPKLNKNEGELTRLRSNIIVTEGFSBEAIEGFOKEIHLNGEERKAKARRKKTLLBDFVA 126
Oy 1224 FIAALVTDQLELVNHFMMVNCFFPRUK--EFLINQOMNDPKSLOQCCLTLRTGSEROI 1281
Db 127 FNGALFLDQGMRAVERFLHLTVPLAKGF--DSRDYKTDLQF-----LQONGPVM 179
Oy 1282 PLYKTLQTPGSHARTYTAAYVFKGERIGCGKPSIQDAEMGAMALEKY 1332
Db 180 EYSVISESGLPSH---FVVELKINDQVQVQEGSHNKKAAEQAKAALQKF 227

RESULT 6
US-11-059-982-1
```

```
; Sequence 1, Application US/11059982
; Publication No. US20050255507A1
; GENERAL INFORMATION:
; APPLICANT: Jenkins, Robert B.
; APPLICANT: Yang, Ping
; APPLICANT: Thibodeau, Steve
; APPLICANT: Wang, Liang
; APPLICANT: Schaid, Daniel
; TITLE OF INVENTION: CYTOGENETICALLY DETERMINED DIAGNOSIS AND
; FILE REFERENCE: 07039-505001
; CURRENT APPLICATION NUMBER: US/11/059,982
; PRIOR FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: US 60/545,573
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-059-982-1

Query Match 2.9%; Score 221; DB 7; Length 1560;
Best Local Similarity 21.7%; Pred. No. 2.4e-06;
Matches 166; Conservative 55; Mismatches 231; Indels 314; Gaps 37;

Oy 4 GNTCHMSFHPGCGCRGNGHARRSAPSPRONRLHLHQPPVQYQYEPSSAPST-- 61
Db 786 GDSFPLPSHPTR-----PPSRPSRQSVS-RPSEPL-HPCPPQAPPTLP 832
Oy 62 ----TFSNS-----PAPNPLPRPDVFPFPPMPPSAQ--GRLPCPIRPPFNH----- 105
Db 833 GIFVIONQGVPRPAPNAPRTAG--PQGPLRQGPREGRLPAPNPLPSSSSAVALS 890
Oy 106 ----QNRHPPVPPCCPRMPPMPCPNPPVPCAPGCGTFPPMPPSPMPPPP-- 157
Db 891 SSETSRLLPAPTSDFQLQFPQCGPKSPTP--PRTL---HLVDEPAPAPPPPPPTFQ 944
Oy 158 ----PVMPQQ----- 167
Db 945 MVTTPPRLPQPKALLERHNVQPSGIILONKAGAPABQTSGLPRLSPASVLSGQ 1004
Oy 168 YPPGY----SHNFP-----PPSFGNFQNNPSSFLPSAN-----NSSPHRHLPYP 211
Db 1005 ABGCTTAPSHAPAPMAATGRLPLPANKAFANLPTLVAKAASSGPKGSLQYE 1064
Oy 212 ----LPKAPSERRSPER--LKHYYDHR---HRDHSQG---RGERHRSIDRR-ERGS 255
Db 1065 SKLSGLKRPPTLOPSKEACFLBHLHKGQSVLHPDYKTAFFSPEDALHRLPYHVOGAL 1124
Oy 256 PRRRRDSDRYRDYDGRTPSRHSYERERERENHRR-----DNR-SPLE----- 304
Db 1125 PS-----PSDYHNKDEEFETVSTQLKKTQAMLNKTKRLLLBESRVSPPAEVMD 1176
Oy 305 ----RKYKEVGRSGRSGVLSVPPACCTPPLPEIINKTDSWAP 346
Db 1177 RMFIOEKTTLLDKOLAKEXDEYVSSRSGLPLAASSEG--HNLPGH--GPLSSAP 1232
Oy 347 ----PLEIYNHRSPPREKKARWEEDKDRNSDQSGKDNVYSIKEPEETW 396
Db 1233 GASTQPPHLPLTFLVIRHGAGGSPVTVARASSLSSSSSSSAASLDADEGCP---M 1289
Oy 397 PDKNF-----EEBELKPYW----- 412
Db 1290 PSNRNPPIKTYEASRIGLKIKOEAAGLSKVHNHTALDPVHOPPPATLKVAPPPRP 1349
Oy 413 ----IRTSSENYSSDPMDOVG----- 431
Db 1350 PPPPPPTGOMGVNHPAPAEKRLGTAHPCRLPLKTYREN-----VGGPCA 1400
Oy 432 -DSTVVGTSR-----LRDLYDKFEELGSRQEK-----AKAR-- 463
```

Db 1401 PEGTPAGRAGSGPAPLPKAVDEATSGILREL-AAVEDELYORMLKGPPEPPASAAGCT 1459
Qy 464 --PPEPPKTKLDEDESSSESCSEDESDSTCSSSSSEVPVIAE 507
Db 1460 GPPDWEAP-----GLPPAKRRKSESPPDVQASFFSSDSFQDDTLTE 1499

RESULT 7

US-11-150-845-38
; Sequence 38, Application US/11150845
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokineics, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-007200US
; CURRENT APPLICATION NUMBER: US/11/150,845
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(559)
; OTHER INFORMATION: SCAR1/WAVE1
US-11-150-845-38

Query Match 2.8%; Score 211.5; DB 7; Length 559;
Best Local Similarity 25.7%; Pred. No. 2.7e-06;
Matches 86; Conservative 30; Mismatches 111; Indels 107; Gaps 16;

Qy 35 RPONTRLHPQOPPVQGYEPP-----SAPSTTFNSNP-----APNFL 72
Db 266 RAERVLVRRHPP-----PPPMHAGDAKPIPTCISATGLIENRPGSPATGRTPV 320
Qy 73 PPRPDVPPPPMPPSA-----OGPLPCCIRPPFNHQRHPPVPCFP--- 118
Db 321 SPTP---PPPPPLPSALSTSLRASMTSTPPPPVPVPPPPPPATLQAPAVPPPAI 377
Qy 119 -----PMPPMPCP---NNPPVGAAPGQGTFFPMMP-----SMPPPPVVPV 164
Db 378 AEGVLHAPAPPIAPLVQSPAPARAPVCETVP-VHPLQGEVQGLPPPPPPPLP--- 433
Qy 165 NYQPPGYSHNPPPSFQNNPSSFLPSANSSSPHFRLLPYPP---LPKAPSERR 220
Db 434 ----PPGIRPSS--PVTYALAHPPSGLHPTSTAGPHVLPMPAPSPQVTP-ASEPKR 486
Qy 221 SPERLKHVDHHRDHSRGGRHRSLDREGRGSPDRRQDSRYSDYDRGRTPSRRHS 280
Db 487 HSTLPLVISD-----ARSVLEAIRKG-----IQLRK 513
Qy 281 YERSRERERHRHNRSPSLERSYKKEYKS 314
Db 514 VEEQREOE-AKHERIENDVATILSRRIAYEYSDS 546

RESULT 8

US-11-149-945-3
; Sequence 3, Application US/11149945
; Publication No. US20060014266A1
; GENERAL INFORMATION:
; APPLICANT: Cytokineics, Inc.

; APPLICANT: Tomasevic, Nenad
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: PURIFICATION OF ARP2/3 COMPLEX AND COMPOSITIONS CONTAINING
; TITLE OF INVENTION: PURIFIED ARP2/3 COMPLEX
; FILE REFERENCE: 020552-007810US
; CURRENT APPLICATION NUMBER: US/11/149,945
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/578,969
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-149-945-3

Query Match 2.8%; Score 211.5; DB 7; Length 559;
Best Local Similarity 25.7%; Pred. No. 2.7e-06;
Matches 86; Conservative 30; Mismatches 111; Indels 107; Gaps 16;

Qy 35 RPONTRLHPQOPPVQGYEPP-----SAPSTTFNSNP-----APNFL 72
Db 266 RAERVLVRRHPP-----PPPMHAGDAKPIPTCISATGLIENRPGSPATGRTPV 320
Qy 73 PPRPDVPPPPMPPSA-----OGPLPCCIRPPFNHQRHPPVPCFP--- 118
Db 321 SPTP---PPPPPLPSALSTSLRASMTSTPPPPVPVPPPPPPATLQAPAVPPPAI 377
Qy 119 -----PMPPMPCP---NNPPVGAAPGQGTFFPMMP-----SMPPPPVVPV 164
Db 378 AEGVLHAPAPPIAPLVQSPAPARAPVCETVP-VHPLQGEVQGLPPPPPPPLP--- 433
Qy 165 NYQPPGYSHNPPPSFQNNPSSFLPSANSSSPHFRLLPYPP---LPKAPSERR 220
Db 434 ----PPGIRPSS--PVTYALAHPPSGLHPTSTAGPHVLPMPAPSPQVTP-ASEPKR 486
Qy 221 SPERLKHVDHHRDHSRGGRHRSLDREGRGSPDRRQDSRYSDYDRGRTPSRRHS 280
Db 487 HSTLPLVISD-----ARSVLEAIRKG-----IQLRK 513
Qy 281 YERSRERERHRHNRSPSLERSYKKEYKS 314
Db 514 VEEQREOE-AKHERIENDVATILSRRIAYEYSDS 546

RESULT 9

US-10-467-657-2638
; Sequence 2638, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: PONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMin99, version 1.04
; SEQ ID NO 2638
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2638

Query Match 2.8%; Score 211; DB 6; Length 244;
Best Local Similarity 28.3%; Pred. No. 1e-06;
Matches 68; Conservative 41; Mismatches 117; Indels 14; Gaps 6;


```

; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-134-563-8

Query Match      2.6%; Score 194.5; DB 7; Length 501;
Best Local Similarity 32.4%; Pred. No. 2,8e-05;
Matches 60; Conservative 12; Mismatches 76; Indels 37; Gaps 7

Oy      82 PPPPPPSAGLPRLPCIRPPFFNQHMRPPRVPRCPMPMPCPN--NPVPGAPRQ 139
Db      274 PPPPPPSRSGPPPP-----PPPHSSGPPPPRPGRGAPPPPSRAPTAAPPPSPRS 329
Oy      140 GFFPPMPRPSMHPHPPPPPVMMQVVVYQVPCGYSHNPPRPSFNSFQNNSSFLLPSANN 199
Db      330 VVP-----PPFNMYPPPPPALPSAPSGPP-----DPL-----SMAGS 366
Oy      200 SPPFRHLPPVPLPKAPSERSRPERLKHYPDHRHRDHSHGREGRRKSLDRREGSPDR 259
Db      367 TAP-----PPPPPPPPPPPPPPPPGLPSGDHGVPASS---GNKALLDQIGEGQLKKV 418
Oy      260 RQDSR 264
Db      419 EONSR 423

RESULT 13
US-11-150-845-16
; Sequence 16, Application US/11150845
; Publication No. US2006000339A1
GENERAL INFORMATION:
; APPLICANT: Cytokinetice, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
FILE REFERENCE: 020552-007720US
CURRENT APPLICATION NUMBER: US/11/150,845
PRIOR FILING DATE: 2005-06-10
PRIOR APPLICATION NUMBER: US 60/673,444
PRIOR FILING DATE: 2005-04-20
PRIOR APPLICATION NUMBER: 60/578,949
PRIOR FILING DATE: 2004-06-10
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.3
SEQ ID NO 16
LENGTH: 716
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Myc-N-WASP-TAP fusion protein
US-11-150-845-16

Query Match      2.6%; Score 192.5; DB 7; Length 716;
Best Local Similarity 21.3%; Pred. No. 5,9e-05;
Matches 94; Conservative 46; Mismatches 174; Indels 127; Gaps 15

Oy      39 LRLHLHQGPVQYQVPPSPASTTFSNAPNPLPRPQFVFPFPPMPSAGLPRLPCPI 98
Db      294 LKRQAPPPPPPSGSGPPPPPPPHSSGPPPP--PARGAGAPPPPSRAPTAAPPPSPRS 350
Oy      99 R-----PPFPNQHMRPPRVPRCPMPMPMPCPN-----NPVPGAPRQGTFFPMMP 148
Db      351 RSVSEVPPPPPRNMYPPPPPALPSAPSGPPPPPSVLGVGVAPRPP-----DP 401
Oy      149 PSMHPDPPPPVMMQVVVYQVPCGYSHNPPRPSFNSFQNNSSFLLPSANNSSPHRHL 208

```

```

Db 402 PPPPPPPPPGLPSDDDHQVF-----TTAGNKAALL-----4320
Qy 209 PVPPLPAMBERSRPRLKTYDDHRRHDSHR-----GEHRSLDRRERGSPD--257
433 -----DQIREGAQLKKVEONSRPVSCSGRDALDQIRGQIQLKSVADQESTPTPA484
Db 258 -----RRRDSRYRDYDRGRTSPRRHSYRERER-----REHRHRDRRS300
Qy 485 PPSIGVIGALMEWQKSKAIHSSDDEDE--DDEDEPDDDEWEDLETMEGRKWKFIYA542
Db 301 PSLERSYKTEKYSRSRYGLSVPEPACCTELPGEIINKTDSWAPLEIYNHRSPSREK360
Qy 543 VSNARFKKIISSGALDLD--LPTTASENLYFQGL--KTAALAQHDEAVDNYF--WKEQ596
Db 361 KRAAME-----BEKDRMSNOSSGCDKNTSYIKEX390
Qy 597 QNPFYELIHLPMINEORNAPIQSLKODPSQSANILAEAKKLTADAQA PKVDNKNENKQON656
Db 391 EPEET--MPDKNEEBEELLK409
Qy 657 AFYEIHLPLNLEEQNAPFIQ677
Db

```

```

RESULT 14
US-11-150-845-14
: Sequence 14, Application US/11150845
: Publication No. US2006003399A1
: GENERAL INFORMATION:
: APPLICANT: Cytokinetics, Inc.
: APPLICANT: Tomasevic, Nenad
: APPLICANT: Jia, Zhiheng
: APPLICANT: Sakowicz, Roman
: APPLICANT: Pierce, Daniel
: APPLICANT: Finer, Jeffrey
: TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
: FILE REFERENCE: 020552-00720US
: CURRENT APPLICATION NUMBER: US/11/150,845
: CURRENT FILING DATE: 2005-06-10
: PRIOR APPLICATION NUMBER: US 60/673,444
: PRIOR FILING DATE: 2005-04-20
: PRIOR APPLICATION NUMBER: 60/578,949
: PRIOR FILING DATE: 2004-06-10
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 14
: LENGTH: 715
: TYPE: PRT
: ORGANISM: Artificial
: FEATURE:
: OTHER INFORMATION: Myc-WASP-TAP fusion protein
US-11-150-845-14

Query Match 2.5%; Score 191; DB 7; Length 715;
Best Local Similarity 21.8%; Pred. No. 7,3e-05;
Matches 90; Conservative 41; Mismatches 130; Indels 152; Gaps 16,

QY 80 PPPPPMPPS-----AOGPLPPCP--TRPPPNKQMHPPPPVPCFPFM 120
    |||||
DB 333 PLPPPPPPSGNQLPPRPVIGNKGKSGPLRPVPLGIAPPPT--PRGPPPPRGSGPP 390
QY 121 PPPMPCNNPPVPAGAPGCGTFPPMMPSPMHPPPPPVMMQVNVYQYPCGVSHNFP 180
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 391 PPPPATGSGPLPPPPGAG-----GPPMPPPPP-----PPF 424
QY 181 SFGSQNNPSSFLPSANSSSPHFRHLPPVPLPRAPSERRRSPERLKHVDHHRDHSQR 240
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 425 -----PSSGNPAP-----PPLPALVPAGGLAP-----G 449
QY 241 GERHSLDRRG-----RSPD-----RRRDSRYPSDYR 271
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 450 GGRGALLDQIRIGIQLNKTFCAPSSALQPPQSSGGLVGLALHMVQKRSRAIHSSDEGE 509

```

QY 272 GTPSRHRSYERSRER--EREPRHRDRRSPSLERSYKKEYKSGRSGYGLSVPEPAG 328
DB 510 DQAGDEDEDEWDEPLETEMERKWKFIIVASANRPKXISSGALDYD--IPTAS 566
QY 329 CTPELPGELIKNTDSWAPLEIVNHRSPREKKARWE----- 366
DB 567 ENLYFOGEL--KTAALAQHDEAVDNKF-NKEQONAFYEILHLPNLNEQORNAFIOSLKDD 623
QY 367 -----BEKDRWSDNOSGKDKNYTSIKEKEPEET--MPDKNEEEELLK 409
DB 624 PQGSANLAEAKKLNDAQAPKYDNKFNKQONAFYEILHLPNLNEQORNAFIQ 676

RESULT 15

US-11-150-845-12
; Sequence 12, Application US/11150845
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokine, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-00720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 408
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-150-845-12

Query Match 2.5%; Score 190.5; DB 7; Length 408;
Best Local Similarity 34.4%; Pred. No. 3.9e-05;
Matches 52; Conservative 9; Mismatches 59; Indels 31; Gaps 7;

QY 19 PRGRGSHGARPSAPSRPQNLRLHPQRPVQYQYERPSASTTSPNSDAPNLPRLPRDF 78
DB 184 PPSRGG---PPPPPPPHN---SGPPPPARGGAGAPPPPSDAPTAAPPP--PPSRPS 234
QY 79 VFPPPPPSAQGLPRLPCIRPPFPNHNQMRHPPRVPPCFPPMPRPPMPCFNNPPVPGAPPG 138
DB 235 VAVPPP-PNNMYPPPP---PALPSS-----APSGPPPPPSVLGVGVAAPPPPP- 280
QY 139 OGTFPPMPPSMHPPPPPVMPQOVNYQY 169
DB 281 -----PPPPPPGPPPPGLPSSDGDHQP 303

Search completed: February 3, 2006, 19:46:56
Job time : 17 secs

THIS PAGE BLANK (USPTO)